

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 18:46:00 ; Search time 1971.58 seconds  
(without alignments)  
9026.108 Million cell updates/sec

Title: US-08-836-455-1

Perfect score: 435

Sequence: 1 ATGGGGGCCCTGCTCAGAT.....CACCATCCAGTAAGCTTGGG 435

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pi.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	435	100.0	435	6	AR164505	AR164505 Sequence
2	435	100.0	435	6	BD085737	BD085737 Methods o
3	377	86.7	381	10	AF124721	AF124721 Mus muscu
4	117	26.9	321	10	AF163749	AF163749 Mus muscu
5	108	24.8	303	10	MUSIGLAF	M36246 Mouse Ig ka
6	108	24.8	324	6	AX722008	AX722008 Sequence
7	107	24.6	302	10	MMU55588	U55588 Mus musculu
8	107	24.6	756	6	AX256284	AX256284 Sequence
9	107	24.6	771	6	AX256296	AX256296 Sequence
10	107	24.6	1497	6	AX256302	AX256302 Sequence
11	102	23.4	387	6	AR169918	AR169918 Sequence
12	100	23.0	269	10	MMVJIG7	X54760 Mouse rearr
13	100	23.0	273	10	MMU21066	U21066 Mus musculu
14	100	23.0	276	6	AR026090	AR026090 Sequence
15	100	23.0	276	6	AR026094	AR026094 Sequence
16	100	23.0	276	10	MMVJIG12	X54765 Mouse rearr
17	100	23.0	279	10	MMVKMRB11	X63811 M.musculus
18	100	23.0	283	10	AF144956	AF144956 Mus muscu
19	100	23.0	285	10	MMU29617	U29617 Mus musculu
20	100	23.0	323	10	AY229938	AY229938 Mus muscu
21	100	23.0	324	6	I03643	I03643 Sequence 4
22	100	23.0	324	6	I07835	I07835 Sequence 4
23	100	23.0	326	10	MMVJIG3	X54756 Mouse rearr
24	100	23.0	348	10	MMVJIG	X54755 Mouse rearr
25	100	23.0	354	10	AB089681	AB089681 Mus muscu
26	100	23.0	381	10	AF045495	AF045495 Mus muscu
27	100	23.0	381	10	AF045508	AF045508 Mus muscu
28	100	23.0	443	10	MMU88675	U88675 Mus musculu
29	100	23.0	684	10	MUSIGKAC1	J00565 Mus musculu
30	100	23.0	685	10	MMIGK7	V00808 Part of the
31	100	23.0	203877	10	AC122260	AC122260 Mus muscu
32	99	22.8	390	10	MUSIKCC	L41880 Mus musculu
33	97	22.3	640	10	MMU242670	AJ242670 Mus muscu
34	96	22.1	218	10	MMVJIG4	X54757 Mouse rearr
35	93	21.4	535	6	BD021878	BD021878 Humanized
36	91	20.9	452	10	AB050077	AB050077 Mus muscu
37	91	20.9	453	10	AB050084	AB050084 Mus muscu
38	91	20.9	456	10	MUSIGKCNK	M19914 Mouse Ig re
39	90	20.7	438	6	E09035	E09035 cDNA encodi
40	90	20.7	438	6	I32991	I32991 Sequence 45
41	89	20.5	321	10	MUSIGKAA3	M59920 Mouse IG ge
42	89	20.5	324	10	MUSX	L48667 Mus musculu
43	89	20.5	381	10	AF045510	AF045510 Mus muscu
44	88	20.2	290	10	AF144902	AF144902 Mus muscu
45	86	19.8	308	10	AF139248	AF139248 Mus muscu

ALIGNMENTS

RESULT 1	AR164505	AR164505	435 bp	DNA	linear	PAT 17-OCT-2001
LOCUS	Sequence 1	from patent US 6274143.				
DEFINITION	Sequence 1	from patent US 6274143.				
ACCESSION	AR164505					
VERSION	AR164505.1	GI:16237555				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 435)					
AUTHORS	Chatterjee,M. and Foon,K.A.					
TITLE	Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11D10					
JOURNAL	Patent: US 6274143-A 1 14-AUG-2001;					

6/11/98

FEATURES	Location/Qualifiers	FT CDS	peptide	1..435
source	1..435	FT mat	Location/Qualifiers	61.
BASE COUNT	100 a 111 c 102 g 122 t			
ORIGIN				
Query Match	100.0%; Score 435; DB 6; Length 435;			
Best Local Similarity	100.0%; Pred. No. 2.6e-258;			
Matches 435; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 ATGGGGCCCTCGCTCAGATCTCTGGGCTCTCTGTGCTCTGTGTTCCAGGTACAGATGT 60			
Db	1 ATGGGGCCCTCGCTCAGATCTCTGGGCTCTCTGTGCTCTGTGTTCCAGGTACAGATGT 60			
QY	61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGAGTCAGT 120			
Db	61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGAGTCAGT 120			
QY	121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACCTTACATTTGGCTTCAGCAGGAACCA 180			
Db	121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACCTTACATTTGGCTTCAGCAGGAACCA 180			
QY	181 GATGGAACCTTATTAACCCCTGATCTACGCCACATCCAGTTAGGTTCTGGTGTCCCAAA 240			
Db	181 GATGGAACCTTATTAACCCCTGATCTACGCCACATCCAGTTAGGTTCTGGTGTCCCAAA 240			
QY	241 AGGTTTCAGTGGCAGTCTGGGCTGAGATTTATCTCCACATCAGCAGCCTTGTAGTCT 300			
Db	241 AGGTTTCAGTGGCAGTCTGGGCTGAGATTTATCTCCACATCAGCAGCCTTGTAGTCT 300			
QY	301 GAAGATTTGTAGCTTATCTGCTACAAATATGCTTCTCCGTCACAGTTTCGGAGGG 360			
Db	301 GAAGATTTGTAGCTTATCTGCTACAAATATGCTTCTCCGTCACAGTTTCGGAGGG 360			
QY	361 GGGACCAAGCTGGAATAAAGCGGCTGATGCTCCACCACTGATCCATCTTCCCAACCA 420			
Db	361 GGGACCAAGCTGGAATAAAGCGGCTGATGCTCCACCACTGATCCATCTTCCCAACCA 420			
QY	421 TCCAGTAAGCTTGGG 435			
Db	421 TCCAGTAAGCTTGGG 435			
RESULT 2				
BD085737				
LOCUS	BD085737 435 bp DNA linear PAT 27-AUG-2002			
DEFINITION	Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11D10.			
ACCESSION	BD085737			
VERSION	BD085737.1 GI:22631347			
KEYWORDS	JP 2001523269-A/1.			
SOURCE	unidentified			
ORGANISM	unclassified.			
REFERENCE	1 (bases 1 to 435)			
AUTHORS	Chatterjee, M. and Foon, K.A.			
TITLE	Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11D10			
JOURNAL	Patent: JP 2001523269-A 1 20-NOV-2001;			
COMMENT	THE UNIVERSITY OF KENTUCKY RESEARCH FOUNDATION			
	OS Unidentified			
	PN JP 2001523269-A/1			
	PD 20-NOV-2001			
	PF 12-JUN-1998 JP 1999503252			
	PR 13-JUN-1997 US 60/049540, 11-JUN-1998 US 09/096244 PI			
	MALAYA CHATTERJEE, KENNETH A. FOON			
	PC A61K39/395.A61K39/39//C07K16/42			
	CC Strandedness: Single;			
	CC Topology: Linear;			
	CC Methods of delaying development of HMFG-associated tumors CC			
	using			
	CC anti-idiotypic antibody 11D10			
	FH Key Location/Qualifiers			

  

FEATURES	Location/Qualifiers	FT CDS	peptide	1..435
source	1..435	FT mat	Location/Qualifiers	61.
BASE COUNT	100 a 111 c 102 g 122 t			
ORIGIN				
Query Match	100.0%; Score 435; DB 6; Length 435;			
Best Local Similarity	100.0%; Pred. No. 2.6e-258;			
Matches 435; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 ATGGGGCCCTCGCTCAGATCTCTGGGCTCTCTGTGCTCTGTGTTCCAGGTACAGATGT 60			
Db	1 ATGGGGCCCTCGCTCAGATCTCTGGGCTCTCTGTGCTCTGTGTTCCAGGTACAGATGT 60			
QY	61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGAGTCAGT 120			
Db	61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGAGTCAGT 120			
QY	121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACCTTACATTTGGCTTCAGCAGGAACCA 180			
Db	121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACCTTACATTTGGCTTCAGCAGGAACCA 180			
QY	181 GATGGAACCTTATTAACCCCTGATCTACGCCACATCCAGTTAGGTTCTGGTGTCCCAAA 240			
Db	181 GATGGAACCTTATTAACCCCTGATCTACGCCACATCCAGTTAGGTTCTGGTGTCCCAAA 240			
QY	241 AGGTTTCAGTGGCAGTCTGGGCTGAGATTTATCTCCACATCAGCAGCCTTGTAGTCT 300			
Db	241 AGGTTTCAGTGGCAGTCTGGGCTGAGATTTATCTCCACATCAGCAGCCTTGTAGTCT 300			
QY	301 GAAGATTTGTAGCTTATCTGCTACAAATATGCTTCTCCGTCACAGTTTCGGAGGG 360			
Db	301 GAAGATTTGTAGCTTATCTGCTACAAATATGCTTCTCCGTCACAGTTTCGGAGGG 360			
QY	361 GGGACCAAGCTGGAATAAAGCGGCTGATGCTCCACCACTGATCCATCTTCCCAACCA 420			
Db	361 GGGACCAAGCTGGAATAAAGCGGCTGATGCTCCACCACTGATCCATCTTCCCAACCA 420			
QY	421 TCCAGTAAGCTTGGG 435			
Db	421 TCCAGTAAGCTTGGG 435			
RESULT 3				
AF124721				
LOCUS	AF124721 381 bp mRNA linear ROD 22-MAY-2001			
DEFINITION	Mus musculus immunoglobulin light chain mRNA, partial cds.			
ACCESSION	AF124721			
VERSION	AF124721.1 GI:14164546			
KEYWORDS				
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	1 (bases 1 to 381)			
AUTHORS	Chatterjee, S.K. and Tripathi, P.K.			
TITLE	Construction and characterization of a chimeric fusion protein consisting of an anti-idiotypic antibody mimicking a breast cancer-associated antigen and the cytokine GM-CSF			
JOURNAL	Hybridoma 18 (2), 193-202 (1999)			
MEDLINE	9305687			
PUBMED	10380019			
REFERENCE	2 (bases 1 to 381)			
AUTHORS	Chatterjee, S.K. and Tripathi, P.K.			
TITLE	Direct Submission			
JOURNAL	Submitted (29-JAN-1999) Internal Medicine, University of Kentucky,			
FEATURES	800 Rose Street, Lexington, KY 40536, USA			
	Location/Qualifiers			



JOURNAL	Submitted (03-JUN-1999) Microbiology and Immunology, Oklahoma University Health Sciences Center, 940 St. Young Blvd, Oklahoma City, OK 73190, USA
FEATURES	Location/Qualifiers
SOURCE	1. 321
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="BALB/c"
	/db_xref="taxon:10090"
	/note="mab 49.8.9"
CDS	<1..>321
	/note="V-J segments"
	/codon_start=1
	/product="immunoglobulin light chain variable region"
	/protein_id="AAD47028.1"
	/db_xref="GI:5690305"
	/translation="DIOMTQSPSSLSASLGERVSLTCRASQDIGSSLNWLQOBDGTI KRLIYATSSLDGVPKRFSGRSRGSYSLTISSEDFVDYVCLQYASSPYFGGGT KLEIK"
BASE COUNT	84 a 77 c 75 g 85 t
ORIGIN	
Query Match	26.9%; Score 117; DB 10; Length 321;
Best Local Similarity	99.1%; Pred. No. 1.5e-60;
Matches	217; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	163 TGGCTTCAGCAGAACACAGATGGAACACTATTAAAGCCCTGATCTACGCCACATCCAGTTTA 222
DB	103 TGGCTTCAGCAGAACACAGATGGAACACTATTAAAGCCCTGATCTACGCCACATCCAGTTTA 162
QY	223 GGTTCCTGGTGCCCCCAAAAGGTTTCAGTGGCAGTAGGCTCTGGTGCAGATTATCTCTCACC 282
DB	163 GATTCTGGTGCCCCCAAAAGGTTTCAGTGGCAGTAGGCTCTGGTGCAGATTATCTCTCACC 222
QY	283 ATCAGCAGCTTGAGCTGGAAGATTTTCTAGCCTATATCTCTACAAATATGCTAGTTCT 342
DB	223 ATCAGCAGCTTGAGCTGGAAGATTTTCTAGCCTATATCTCTACAAATATGCTAGTTCT 282
QY	343 CGGTACACGTTTCGGAGGGGGACCAAGCTGGAAATAAAA 381
DB	283 CGGTACACGTTTCGGAGGGGGACCAAGCTGGAAATAAAA 321
RESULT 5	
LOCUS	MUSIGLAF 303 bp mRNA linear ROD 27-APR-1993
DEFINITION	Mouse Ig kappa-chain mRNA V region, partial cds, from hybridoma H220-23.
ACCESSION	M36246
VERSION	M36246.1 GI:197631
KEYWORDS	V-region; immunoglobulin kappa-chain; immunoglobulin light chain; processed gene.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 303)
TITLE	Kavaler, J., Caton, A.J., Staudt, L.M., Schwartz, D. and Gerhard, W. A set of closely related antibodies dominates the primary antibody response to the antigenic site CB of the A/PR/8/34 influenza virus hemagglutinin
JOURNAL	J. Immunol. 145 (7), 2312-2321 (1990)
MEDLINE	90375932
PUBMED	2118933
COMMENT	Original source text: Mouse (strain BALB/c), CDNA to mRNA, from hybridoma H220-23. Draft entry and computer-readable sequence for [J. Immunol. (1990) In press] kindly submitted by J.Kavaler, 06-JUL-1990.
FEATURES	Location/Qualifiers
SOURCE	1. 303
	/organism="Mus musculus"
	/mol_type="mRNA"

QY	285	CACAGCCTTGAGTCTGAAGATTTTGTAGCCTATTACTGCTACAATATGCTAGTTCTCC	344
Db	225	CACAGCCTTGAGTCTGAAGATTTTGTAGCCTATTACTGCTACAATATGCTAGTTCTCC	284
QY	345	GTACAGCTTCGGAGGGGGGACCAAGCTGGAAATAAAACG	383
Db	285	GTACAGCTTCGGAGGGGGGACCAAGCTGGAAATAAAACG	323
RESULT 7			
MMU55588			
LOCUS		302 bp mRNA linear	ROD 14-SEP-2001
DEFINITION		Mus musculus anti-DNA immunoglobulin light chain IgG, antibody	
ACCESSION		U55588	363s-62, partial cds.
VERSION		U55588.1	GI:1870291
KEYWORDS			
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
TITLE		1 (bases 1 to 302)	
JOURNAL		Krishnan,M.R., Jou,N.T. and Marion,T.N.	
MEDLINE		Correlation between the amino acid position of arginine in VH-CDR3	
PUBMED		and specificity for native DNA among autoimmune antibodies	
REFERENCE		J. Immunol. 157 (6), 2430-2439 (1996)	
AUTHORS		96399071	
TITLE		2 (bases 1 to 302)	
JOURNAL		Marion,T.N.	
MEDLINE		Direct Submission	
PUBMED		Submitted (18-APR-1996) Tony N. Marion, Dept. of	
AUTHORS		Microbiology/Immunology, University of Tennessee, 858 Madison Ave.,	
TITLE		Memphis, TN 38163, USA	
JOURNAL		1. .302	
MEDLINE		Location/Qualifiers	
PUBMED		/organism="Mus musculus"	
AUTHORS		/mol_type="mRNA"	
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
AUTHORS			

```

/db_xref=taxon:10090
/cell_line=363s.62"
/cell_type="hybridoma"
/note="mouse number 363"
<1..>302
/codon_start=1
/product="anti-DNA immunoglobulin light chain IgG"
/protein_id="AAB48706.1"
/db_xref="GI:1870292"
/translation="MTXSPSSLSASLGERVSLTCRASQDITGSSLNLKQQEPDGTIKRL
IVATSLDSGVPKRFSGSRGSVDYSLTISSELEDFDYCYLQYASSPTFFGGTKL"
V_region      74 a    72 c    71 g    84 t    1 others
BASE COUNT   <1..>302
ORIGIN
Query Match          24.6%; Score 107; DB 10; Length 302;
Best Local Similarity 99.0%; Pred. No. 2.4e-54;
Matches 207; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      163 TGCGTTTCAGCAGGAACCCAGATGGAACCTATTAAACGCCTGATCTACGCCATCATCCAGTTTA 222
           |
Db       94 TGCGTTTCAGCAGGAACCCAGATGGAACCTATTAAACGCCTGATCTACGCCATCATCCAGTTTA 153
           |||
QY      223 GGTTCTGGTGTCCTCCCAAAAGGTTTCAGTGGCAGTAGCTCTGGGTCAGATTATTTCTCTCAC 282
           |
Db     154 GATTCTGGTGTCCTCCCAAAAGGTTTCAGTGGCAGTAGCTCTGGGTCAGATTATTTCTCTCAC 213
           |||
QY      283 ATCAGCAGCCTTCAGTCTGAAGATTTTGTAGCCCTATTACTGTCTACAATATGCTAGTTCT 342
           |||
Db     214 ATCAGCAGCCTTCAGTCTGNAGATTTTGTAGACTATTACTGTCTACAATATGCTAGTTCT 273
           |||
QY      343 CCGTACACGTTTCGGAGGGGGGACCACCAAGCT 371
           |||

```

Query Match	24.6%;	Score 107;	DB 6;	Length 771;					
Best Local Similarity	99.0%;	Pred. No. 2.5e-54;							
Matches	207;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	163	TGCGCTTCAGCAGAACCAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	222						
Db	517	TGCGCTTCAGCAGAACCAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	576						
QY	223	GCTTCTGGTGTGTCCTCCCAAAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	282						
Db	577	GATTCTGGTGTGTCCTCCCAAAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	636						
QY	283	ATCAGCAGCTTGAGTCTGAAGATTTTGTACCCCTATTACTGCTACAAATATGCTAGTTCT	342						
Db	637	ATCAGCAGCTTGAGTCTGAAGATTTTGTACCCCTATTACTGCTACAAATATGCTAGTTCT	696						
QY	343	CCGTACAGCTTCGGAGGGGGACCAAGCT	371						
Db	697	CCGTACAGCTTCGGAGGGGGACCAAGCT	725						
RESULT 10									
AX256302									
LOCUS	AX256302		1497 bp	DNA	linear	PAT 10-OCT-2001			
DEFINITION	Sequence 72 from Patent WO0171005.								
ACCESSION	AX256302								
VERSION	AX256302.1	GI:16075169							
KEYWORDS	synthetic construct								
SOURCE	synthetic construct								
ORGANISM	artificial sequences.								
REFERENCE	1								
AUTHORS	Kufer, P., Riethmüller, G., Lutterbuese, R., Borschert, K., Kischel, R., Mayer, M. and Hofmeister, R.								
TITLE	Multifunctional polypeptides comprising a binding site to an epitope of the nk92d receptor complex								
JOURNAL	Patent: WO 0171005-A 72 27-SEP-2001;								
FEATURES	Kufer, Peter (DE)								
source	Location/Qualifiers								
	1..1497								
	/organism="synthetic construct"								
	/mol_type="genomic DNA"								
	/db_xref="taxon:32630"								
BASE COUNT	364 a	358 c	407 g	368 t					
ORIGIN									
Query Match	24.6%;	Score 107;	DB 6;	Length 1497;					
Best Local Similarity	99.0%;	Pred. No. 2.5e-54;							
Matches	207;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	163	TGCGCTTCAGCAGAACCAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	222						
Db	1243	TGCGCTTCAGCAGAACCAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	1302						
QY	223	GCTTCTGGTGTGTCCTCCCAAAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	282						
Db	1303	GATTCTGGTGTGTCCTCCCAAAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	1362						
QY	283	ATCAGCAGCTTGAGTCTGAAGATTTTGTACCCCTATTACTGCTACAAATATGCTAGTTCT	342						
Db	1363	ATCAGCAGCTTGAGTCTGAAGATTTTGTACCCCTATTACTGCTACAAATATGCTAGTTCT	1422						
QY	343	CCGTACAGCTTCGGAGGGGGACCAAGCT	371						
Db	1423	CCGTACAGCTTCGGAGGGGGACCAAGCT	1451						
RESULT 11									
AX169918									
LOCUS	AX169918		387 bp	DNA	linear	PAT 17-DEC-2001			
DEFINITION	Sequence 1 from patent US 6291208.								
ACCESSION	AR169918								
VERSION	AR169918.1	GI:17907877							
QY	163	TGCGCTTCAGCAGAACCAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	222						
Db	517	TGCGCTTCAGCAGAACCAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	576						
QY	223	GCTTCTGGTGTGTCCTCCCAAAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	282						
Db	577	GATTCTGGTGTGTCCTCCCAAAAGGTTTCAGTGGCAGTGGGCTGAGTCTACGCCACATCCAGTTTA	636						
QY	283	ATCAGCAGCTTGAGTCTGAAGATTTTGTACCCCTATTACTGCTACAAATATGCTAGTTCT	342						
Db	637	ATCAGCAGCTTGAGTCTGAAGATTTTGTACCCCTATTACTGCTACAAATATGCTAGTTCT	696						
QY	343	CCGTACAGCTTCGGAGGGGGACCAAGCT	371						
Db	697	CCGTACAGCTTCGGAGGGGGACCAAGCT	725						
RESULT 10									
AX256302									
LOCUS	AX256302		1497 bp	DNA	linear	PAT 10-OCT-2001			
DEFINITION	Sequence 72 from Patent WO0171005.								
ACCESSION	AX256302								
VERSION	AX256302.1	GI:16075169							
KEYWORDS	synthetic construct								
SOURCE	synthetic construct								
ORGANISM	artificial sequences.								
REFERENCE	1								
AUTHORS	Kufer, P., Riethmüller, G., Lutterbuese, R., Borschert, K., Kischel, R., Mayer, M. and Hofmeister, R.								
TITLE	Multifunctional polypeptides comprising a binding site to an epitope of the nk92d receptor complex								

```
KEYWORDS
SOURCE      Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 387)
AUTHORS      Anand,N.N., Barber,B.H., Cates,G.C., Caterini,J.E. and Klein,M.H.
TITLE        Chimeric antibodies for delivery of antigens to selected cells of
              the immune system
JOURNAL      Patent: US 6291208-A 1 18-SEP-2001;
FEATURES     Location/Qualifiers
              source
              1..387
              /organism="unknown"
BASE COUNT   90 a 95 c 112 t
ORIGIN
Query Match      23.4%; Score 102; DB 6; Length 387;
Best Local Similarity 100.0%; Pred. No. 3.2e-51;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 GTTCCAGGTACACAGATGACATCCAGATGACCCAGTCTCCATCCCTTATCTGCCTC 101
    |||||||
Db 48 GTTCCAGGTACACAGATGACATCCAGATGACCCAGTCTCCATCCCTTATCTGCCTC 107

Qy 102 TCTGGGCAAGAGTCAGTCTCACTTCTCGGGCAAGTCAGGA 143
    |||||||
Db 108 TCTGGGCAAGAGTCAGTCTCACTTCTCGGGCAAGTCAGGA 149

RESULT 12
MMVJIG7
LOCUS      MMVJIG7                269 bp    DNA    linear    ROD 25-JUL-2002
DEFINITION Mouse rearranged kappa immunoglobulin light chain (V.J).
ACCESSION  X54760
VERSION    X54760.1 GI:21998644
KEYWORDS   Ig light chain; immunoglobulin.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 269)
AUTHORS    Harada,K. and Yamagishi,H.
TITLE      Lack of feedback inhibition of V kappa gene rearrangement by
            productively rearranged alleles
JOURNAL    J. Exp. Med. 173 (2), 409-415 (1991)
MEDLINE    91108338
PUBMED     1988542
REFERENCE  2 (bases 1 to 269)
AUTHORS    Yamagishi,H.
TITLE      Direct Submision
JOURNAL    Submitted (02-OCT-1990) Yamagishi H., Dept of Biophysics Faculty of
            Science Kyoto University, Sakyo-ku Kyoto 606, Japan
FEATURES   Location/Qualifiers
            source
            1..269
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /strain="BALB/c-nu/nu"
            /db_xref="taxon:10090"
            /clone="MSI-N105"
            /cell_type="splenocyte"
            /tissue_type="spleen"
            /clone_lib="lambda dtl1"
            /dev_stage="8 week old"
            <1..>269
            /codon_start=2
            /product="Ig light chain"
            /protein_id="CAA38558.1"
            /db_xref="GI:21998645"
            /translation="SLTCRASQDGLSSLNWLQEPDGTIKRLIYATSSLDGVPKRES
            GSRSGSDISLTISSESEDFDYICLOYASSPVDVRRWRHQNGQT"
CDS
mat_peptide
prim_transcript <1..>269
misc_feature    1..199

/Note="V kappa gene segment"
200..267
/Note="J kappa gene segment"
74 t

misc_feature    69 a 61 c 65 g 74 t
ORIGIN
Query Match      23.0%; Score 100; DB 10; Length 269;
Best Local Similarity 99.3%; Pred. No. 5.5e-50;
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 163 TGGCTTCAGCAGGAACACAGATGGAACATATTAAAGCCCTGATCTACGCCACATCCAGTTTA 222
    |||||||
Db 47 TGGCTTCAGCAGGAACACAGATGGAACATATTAAAGCCCTGATCTACGCCACATCCAGTTTA 106

Qy 223 GGTTCGTGGTGTCCCCAAAGGTTTCAGTGGCAGTAGGTCTGGGTACAGATTATTTCTCACC 282
    |||||||
Db 107 GATTCTGGTGTCCCCAAAGGTTTCAGTGGCAGTAGGTCTGGGTACAGATTATTTCTCACC 166

Qy 283 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAG 313
    |||||||
Db 167 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAG 197

RESULT 13
MMU21066
LOCUS      MMU21066                273 bp    mRNA    linear    ROD 14-FEB-1996
DEFINITION Mus musculus immunoglobulin kappa chain V-J regions mRNA, clone
            MRL3-7, partial cds.
ACCESSION  U21066
VERSION    U21066.1 GI:699556
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 273)
AUTHORS    Roark,J.H., Kuntz,C.L., Nguyen,K.A., Mandik,L., Cattermole,M. and
            Erikson,J.
TITLE      B cell selection and allelic exclusion of an anti-DNA Ig transgene
            in MRL-lpr/lpr mice
JOURNAL    J. Immunol. 154 (9), 4444-4455 (1995)
MEDLINE    95238942
PUBMED     7722301
REFERENCE  2 (bases 1 to 273)
AUTHORS    Roark,J.H.
TITLE      Direct Submission
JOURNAL    Submitted (15-FEB-1995) Jessica H. Roark, Wistar Institute, 3601
            Spruce St., Philadelphia, PA 19104, USA
FEATURES   Location/Qualifiers
            source
            1..273
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="MRL-lpr/lpr"
            /db_xref="taxon:10090"
            /clone="MRL3-7"
            /cell_type="splenic B cell hybridoma"
            /tissue_type="spleen"
            /dev_stage="adult"
            <1..>273
            /codon_start=3
            /product="immunoglobulin kappa chain variable and joining
            regions"
            /protein_id="AA87450.1"
            /db_xref="GI:699557"
            /translation="LSGRKKSQHSRASQDGLSSLNWLQEPDGTIKRLIYATSSLDG
            VPKRFGSRSGSDYSLTISSESEDFDYICLOYAFSPYFGSGTK"
BASE COUNT   68 a 63 c 66 g 76 t
ORIGIN
Query Match      23.0%; Score 100; DB 10; Length 273;
Best Local Similarity 99.3%; Pred. No. 5.5e-50;
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 163 TGGCTTCAGCAGAACCCAGATGGAACACTATTAAACGCGCTGATCTAGCCACATCCAGTTTA 222  
|||||  
Db 66 TGGCTTCAGCAGAACCCAGATGGAACACTATTAAACGCGCTGATCTAGCCACATCCAGTTTA 125  
|||||  
QY 223 GGTTCCTGGTGTCCTCCCAAGGTTTCAGTGGCAGTAGGCTCTGGCTCAGATTATTCTCTCACC 282  
|||||  
Db 126 GATTCTGGTGTCCTCCCAAGGTTTCAGTGGCAGTAGGCTCTGGCTCAGATTATTCTCTCACC 185  
|||||  
QY 283 ATCAGCAGCCTTGAGCTCTGAAGATTTTGTAG 313  
|||||  
Db 186 ATCAGCAGCCTTGAGCTCTGAAGATTTTGTAG 216  
|||||

RESULT 14  
AR026090  
LOCUS AR026090 276 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 44 from patent US 5855885.  
ACCESSION AR026090  
VERSION AR026090.1 GI:5936930  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.

REFERENCE 1 (bases 1 to 276)  
AUTHORS Smith,R., McCafferty,J., Chiswell,D., Darsley,M.J., Fitzgerald,K.,  
Kenten,J.H., Martin,M.T., Titmas,R.C. and Williams,R.O.  
TITLE Isolation and production of catalytic antibodies using phage  
technology  
JOURNAL Patent: US 5855885-A 44 05-JAN-1999;  
FEATURES location/Qualifiers  
source l..276  
/organism="unknown"

BASE COUNT 68 a 65 c 63 g 80 t  
ORIGIN

Query Match 23.0%; Score 100; DB 6; Length 276;  
Best Local Similarity 99.3%; Pred. No. 5.5e-50;  
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 163 TGGCTTCAGCAGAACCCAGATGGAACACTATTAAACGCGCTGATCTAGCCACATCCAGTTTA 222  
|||||  
Db 79 TGGCTTCAGCAGAACCCAGATGGAACACTATTAAACGCGCTGATCTAGCCACATCCAGTTTA 138  
|||||  
QY 223 GGTTCCTGGTGTCCTCCCAAGGTTTCAGTGGCAGTAGGCTCTGGCTCAGATTATTCTCTCACC 282  
|||||  
Db 139 GATTCTGGTGTCCTCCCAAGGTTTCAGTGGCAGTAGGCTCTGGCTCAGATTATTCTCTCACC 198  
|||||  
QY 283 ATCAGCAGCCTTGAGCTCTGAAGATTTTGTAG 313  
|||||  
Db 199 ATCAGCAGCCTTGAGCTCTGAAGATTTTGTAG 229  
|||||

RESULT 15  
AR026094  
LOCUS AR026094 276 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 52 from patent US 5855885.  
ACCESSION AR026094  
VERSION AR026094.1 GI:5936934  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.

REFERENCE 1 (bases 1 to 276)  
AUTHORS Smith,R., McCafferty,J., Chiswell,D., Darsley,M.J., Fitzgerald,K.,  
Kenten,J.H., Martin,M.T., Titmas,R.C. and Williams,R.O.  
TITLE Isolation and production of catalytic antibodies using phage  
technology  
JOURNAL Patent: US 5855885-A 52 05-JAN-1999;  
FEATURES location/Qualifiers  
source l..276  
/organism="unknown"

BASE COUNT 67 a 65 c 64 g 80 t  
ORIGIN

Query Match 23.0%; Score 100; DB 6; Length 276;  
Best Local Similarity 99.3%; Pred. No. 5.5e-50;  
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 163 TGGCTTCAGCAGAACCCAGATGGAACACTATTAAACGCGCTGATCTAGCCACATCCAGTTTA 222  
|||||  
Db 79 TGGCTTCAGCAGAACCCAGATGGAACACTATTAAACGCGCTGATCTAGCCACATCCAGTTTA 138  
|||||  
QY 223 GGTTCCTGGTGTCCTCCCAAGGTTTCAGTGGCAGTAGGCTCTGGCTCAGATTATTCTCTCACC 282  
|||||  
Db 139 GATTCTGGTGTCCTCCCAAGGTTTCAGTGGCAGTAGGCTCTGGCTCAGATTATTCTCTCACC 198  
|||||  
QY 283 ATCAGCAGCCTTGAGCTCTGAAGATTTTGTAG 313  
|||||  
Db 199 ATCAGCAGCCTTGAGCTCTGAAGATTTTGTAG 229  
|||||

Search completed: August 30, 2003, 21:04:54  
Job time : 1972.58 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 17:38:10 ; Search time 190.312 Seconds  
(without alignments)  
6170.144 Million cell updates/sec

Title: US-08-836-455-1  
Perfect score: 435  
Sequence: 1 ATGGGGGCCCTGCTCAGT.....CACCATCCAGTAAGCTTGGG 435

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2552736 seqs, 1349719017 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_13Jun03:  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	435	100.0	435	AA785149	Murine monoclonal
2	435	100.0	435	AA783772	Antibody 11D10 lig
3	435	100.0	435	AA151273	Mouse 11D10 antibo
4	107	24.6	756	AA597136	P4-3 single chain
5	107	24.6	771	AA597142	P5-10 single chain
6	107	24.6	1497	AA597145	3B10xP4-3 bispecif
7	102	23.4	387	AA777851	Murine anti-human
8	102	23.4	387	AA778218	Murine 44H104 mab

9	100	23.0	276	20	AA000879	Mouse derived RT3
10	100	23.0	276	20	AA000875	Mouse derived RT3
11	100	23.0	450	4	AA030165	Sequence encoding
12	93	21.4	535	19	AA020086	Consensus DNA sequ
13	90	20.7	438	16	AA090431	DNA encoding antio
14	86	19.8	979	24	AB52791	Murine m166 antibo
15	85	19.5	642	18	AA085091	Human secreted exp
16	85	19.5	651	21	AA044346	Human secreted exp
17	85	19.5	652	17	AA087818	Antibody 3G2 light
18	85	19.5	654	25	AA044908	TSH receptor antib
19	85	19.5	654	25	AA044909	TSH receptor antib
20	85	19.5	678	21	AA027849	WOW-1 Fab light ch
21	85	19.5	698	25	ABX16574	Mouse DNA encoding
22	85	19.5	723	16	AA092503	Mouse antibody F4-
23	85	19.5	738	21	AA061037	Nucleotide sequenc
24	84	19.3	639	10	AA091657	Chimeric antibody
25	83	19.1	465	16	AA085387	MAD 4197X light ch
26	82	18.9	324	21	AA049534	Mouse anti-IL-18 a
27	82	18.9	407	21	AA049548	Mouse light chain
28	82	18.9	711	21	AA049542	PESCFV#125-2H reco
29	82	18.9	729	21	AA049543	PESCFV#125-2H HT r
30	81	18.6	882	14	AA048038	Monoclonal antibody
31	80	18.4	654	25	AA044928	TSH receptor antib
32	80	18.4	654	25	AA044929	TSH receptor antib
33	80	18.4	729	25	ABX16570	Mouse DNA encoding
34	79	18.2	426	22	AA013177	Murine antibody 1D
35	79	18.2	438	18	AA070809	Mouse anti-idiotyp
36	79	18.2	537	12	AA014801	Encodes murine ant
37	79	18.2	537	12	AA014651	R6-5-D6 anti-ICAM-
38	78	17.9	366	13	AA027140	ICAM-1 inhibiting
39	78	17.9	451	21	AA043470	Mouse secreted exp
40	77	17.7	504	25	ABX16572	Mouse DNA encoding
41	73	16.8	321	21	AA038909	650E2 hybridoma VL
42	73	16.8	456	22	AA066996	Filamentous phage
43	73	16.8	867	22	AA067002	Filamentous phage
44	70	16.1	360	18	AA096344	CDNA for Ig light
45	70	16.1	1848	18	AA096346	Chimeric gene cont

## ALIGNMENTS

RESULT 1  
AA785149  
ID AA785149 standard; CDNA; 435 BP.  
XX  
AC AA785149;  
XX  
DT 25-MAR-2003 (updated)  
DT 04-JAN-1998 (first entry)  
XX  
DE Murine monoclonal anti-idiotyp antibody 11D10 VL cDNA.  
XX  
KW Monoclonal antibody 11D10; anti-idiotyp antibody; mucin;  
KW human milk fat globule; HMFG; tumour; breast cancer; vaccine; ss.  
XX  
OS Mus musculus.  
XX  
FH Key  
FT sig\_peptide Location/Qualifiers  
FT mat\_peptide 1..60  
FT 61..435 /\*tag- a  
FT /\*tag- b

WO9722699-A2.

XX  
XX 26-JUN-1997.  
XX  
XX 19-DEC-1996;  
XX 96WO-US20757.  
PR 20-DEC-1995;  
PR 26-JAN-1996;  
PR 13-DEC-1996;  
PR 95US-0575762.  
PR 96US-0591965.  
PR 96US-0766350.





QY 241 AGGTTTCAGTGGCAGTAGTCTGGGTCAGATTATTTCTCTCACCATCAGCAGCCTTGAGTCT 300  
 |||||  
 Db 241 AGGTTTCAGTGGCAGTAGTCTGGGTCAGATTATTTCTCTCACCATCAGCAGCCTTGAGTCT 300  
 |||||  
 QY 301 GAAGATTTTGTAGCCTATTACTGTCTACAATAGTCTCTCCGTACACGTTCCGGAGG 360  
 |||||  
 Db 301 GAAGATTTTGTAGCCTATTACTGTCTACAATAGTCTCTCCGTACACGTTCCGGAGG 360  
 |||||  
 QY 361 GGGACCAAGCTGGGAATTAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
 |||||  
 Db 361 GGGACCAAGCTGGGAATTAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
 |||||  
 QY 421 TCCAGTAAGCTTGGG 435  
 |||||  
 Db 421 TCCAGTAAGCTTGGG 435  
 |||||

## RESULT 3

AAL51273  
 ID AAL51273 standard; cDNA; 435 BP.

AC AAL51273;

DT 20-MAR-2003 (first entry)

DE Mouse l1d10 antibody light chain variable region coding sequence.

KW Mouse; murine; gene; ss; vaccine; tumour; human milk fat globules; HMFG;  
 KW carcinoembryonic antigen; CEA; l1d10; 3H1; HMFG-associated tumour;  
 KW CEA-associated tumour; anti-idiotype antibody.

OS Mus musculus.

FH Key Location/Qualifiers

FT 1..435

FT /\*tag- a

FT /product-

FT "Mouse l1d10 anti-idiotype antibody light chain

FT variable region"

FT /note= "No stop codon is given"

FT sig\_peptide 1..60

FT /\*tag- b

FT mat\_peptide 61..435

FT /\*tag- c

PN WO200292012-A2.

PD 21-NOV-2002.

PF 17-MAY-2002; 2002WO-US15840.

PR 17-MAY-2001; 2001US-0861294.

PA (KENT ) UNIV KENTUCKY RES FOUND.

XX Chatterjee M, Foon KA;

XX WPI: 2003-129216/12.

XX P-PSDB; AA016292.

XX Use of anti-idiotypic antibodies for human milk fat globules (HMFG) - or

XX carcinoembryonic antigen (CEA)-associated tumor for delaying the

XX development of, or treating a HMFG- or CEA-associated tumor (e.g.

XX breast tumor) in humans

XX Disclosure; Fig 1; 98pp; English.

XX The invention comprises a method for delaying the development of, or

XX treating a tumour that is associated with human milk fat globules (HMFG)

XX or carcinoembryonic antigen (CEA). The method of the invention involves

XX administering an anti-idiotypic antibody for HMFG (e.g. l1d10), and an

XX anti-idiotypic antibody for CEA (e.g. 3H1). The method is useful for

CC delaying the development, of or treating HMFG/CEA-associated tumours. The  
 CC present cDNA sequence encodes the light chain variable region of the  
 CC mouse l1d10 anti-idiotype antibody.

SQ Sequence 435 BP; 100 A; 111 C; 102 G; 122 T; 0 other;

Query Match 100.0%; Score 435; DB 25; Length 435;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-210;  
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGGCCCCGCTCAGATTCTTGGGTTCTTGGTCTCTTGTCTTCCAGGTACCAGATGT 60  
 |||||

Db 1 ATGGGGGGCCCCGCTCAGATTCTTGGGTTCTTGGTCTCTTGTCTTCCAGGTACCAGATGT 60  
 |||||

QY 61 GACATCCAGATGACCCAGTCTCCATCCCTCTTATCTGCGCTCTCTGGGACAAAGAGTCAGT 120  
 |||||

Db 61 GACATCCAGATGACCCAGTCTCCATCCCTCTTATCTGCGCTCTCTGGGACAAAGAGTCAGT 120  
 |||||

QY 121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACCTTACATTTGGCTTCAGCAGGAACCA 180  
 |||||

Db 121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACCTTACATTTGGCTTCAGCAGGAACCA 180  
 |||||

QY 181 GATGGAACCTATTAAACGGCTGATCTACGCCACATCCAGTTTAGGTTCTGGTCCCCCAA 240  
 |||||

Db 181 GATGGAACCTATTAAACGGCTGATCTACGCCACATCCAGTTTAGGTTCTGGTCCCCCAA 240  
 |||||

QY 241 AGGTTTCAGTGGCAGTAGTCTGGGTCAGATTATTTCTCTCACCATCAGCAGCCTTGAGTCT 300  
 |||||

Db 241 AGGTTTCAGTGGCAGTAGTCTGGGTCAGATTATTTCTCTCACCATCAGCAGCCTTGAGTCT 300  
 |||||

QY 301 GAAGATTTTGTAGCCTATTACTGTCTACAATATGCTTCCTCCGTCACGTTCCGGAGG 360  
 |||||

Db 301 GAAGATTTTGTAGCCTATTACTGTCTACAATATGCTTCCTCCGTCACGTTCCGGAGG 360  
 |||||

QY 361 GGGACCAAGCTGGGAATTAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
 |||||

Db 361 GGGACCAAGCTGGGAATTAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
 |||||

QY 421 TCCAGTAAGCTTGGG 435

Db 421 TCCAGTAAGCTTGGG 435

## RESULT 4

AAS97136

ID AAS97136 standard; DNA; 756 BP.

XX AAS97136;

AC AAS97136;

DT 26-FEB-2002 (first entry)

DE P4-3 single chain Fv DNA.

XX Human; NKG2D; NKG2D receptor complex; cancer; infectious disease; tumour;

XX autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;

XX intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;

XX prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;

XX sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; ds;

XX helminth; cytostatic; antimicrobial; immunomodulatory; P4-2; P4-3; P4-14;

XX P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10XP4-3; 3B10XP4-14;

XX P53 tetramerisation domain; 3B10XP5-2; 3B10XP5-23.

XX Homo sapiens.

OS Homo sapiens.

PN WO200171005-A2.

XX 27-SEP-2001.

XX 26-MAR-2001; 2001WO-EP03414.

XX 24-MAR-2000; 2000EP-0106467.

XX (KUFE/) KUFER P.

```

XX PI Kufer P, Rietmueller G, Lutterbuese R, Borschert K, Kischel R;
XX PI Mayer M, Hofmeisler R;
XX DR WPI; 2002-055119/07.
XX DR P-PSDB; AAU72862.
XX
XX PT Multifunctional polypeptides comprising binding sites that specifically
XX PT recognise extracellular groups of the NKG2D receptor complex and
XX PT domains which function as receptors or ligands, useful for treating
XX PT cancers and infectious diseases -
XX PS
XX PS Example 7; Fig 16; 114pp; English.
XX
XX CC The invention relates to a multifunctional polypeptide comprising a
XX CC domain with a binding site that specifically recognises an extracellular
XX CC group of the NKG2D receptor complex and a second domain which functions
XX CC as a receptor or ligand. The polypeptide and its associated
XX CC polynucleotide are used for the preparation of a pharmaceutical
XX CC composition for the treatment of cancer, infections and/or autoimmune
XX CC conditions. The cancer may be a tumour of the head and neck, stomach,
XX CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,
XX CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,
XX CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.
XX CC The infectious diseases can be caused by viruses, bacteria, fungi,
XX CC protozoa or helminths. The autoimmune diseases include multiple
XX CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior
XX CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent
XX CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and
XX CC autoimmune hepatitis. The sequences represent DNA encoding the
XX CC polypeptides of the invention.
XX
XX SQ Sequence 756 BP; 184 A; 180 C; 204 G; 188 T; 0 other;
Query Match 24.6%; Score 107; DB 24; Length 756;
Best Local Similarity 99.0%; Pred. No. 5.5e-44;
Matches 207; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 163 TGGCTTCAGCAGGAACCCAGATGGAAGTATTTAAACGGCTGATCTACGCCACATCCAGTTTA 222
XX DB 502 TGGCTTCAGCAGGAACCCAGATGGAAGTATTTAAACGGCTGATCTACGCCACATCCAGTTTA 561
XX
XX QY 223 GGTTCCTGGTGTCCCAAAAGGTTTCAGTGGCAGTAGTCTGGGTCAGATTATTTCTTCACC 282
XX DB 562 GATTCCTGGTGTCCCAAAAGGTTTCAGTGGCAGTAGTCTGGGTCAGATTATTTCTTCACC 621
XX
XX QY 283 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAGCCCTATTACTGTCTACAATATGCTAGTTCT 342
XX DB 622 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAGACTATTACTGTCTACAATATGCTAGTTCT 681
XX
XX QY 343 CCGTACACGTTTCGGAGGGGGGACCAAGCT 371
XX DB 682 CCGTACACGTTTCGGAGGGGGGACCAAGCT 710
XX
XX RESULT 5
XX ID AAS97142
XX AC AAS97142 standard; DNA; 771 BP.
XX
XX AC AAS97142;
XX
XX DT 26-FEB-2002 (first entry)
XX
XX DE p5-10 single chain Fv DNA.
XX
XX KW Human; NKG2D; NKG2D receptor complex; cancer; infectious disease; tumour;
XX KW autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;
XX KW intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;
XX KW prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;
XX KW sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; ds;
XX KW helminth; cytostatic; antimicrobial; immunomodulatory; p4-2; p4-3; p4-14;
XX KW p4-15; p5-2; p5-3; p5-9; p5-10; p5-11; p5-23; 3B10xp4-3; 3B10xp4-14;
XX KW p53 tetramerisation domain; 3B10xp5-2; 3B10xp5-23.

```

```

XX OS Homo sapiens.
XX PN WO200171005-A2.
XX XX 27-SEP-2001.
XX
XX XX 26-MAR-2001; 2001WO-EP03414.
XX XX 24-MAR-2000; 2000EP-0106467.
XX XX (KUFE/) KUFER P.
XX
XX PI Kufer P, Rietmueller G, Lutterbuese R, Borschert K, Kischel R;
XX PI Mayer M, Hofmeisler R;
XX XX WPI; 2002-055119/07.
XX XX P-PSDB; AAU72868.
XX
XX PT Multifunctional polypeptides comprising binding sites that specifically
XX PT recognise extracellular groups of the NKG2D receptor complex and
XX PT domains which function as receptors or ligands, useful for treating
XX PT cancers and infectious diseases -
XX PS
XX PS Example 7; Fig 16; 114pp; English.
XX
XX CC The invention relates to a multifunctional polypeptide comprising a
XX CC domain with a binding site that specifically recognises an extracellular
XX CC group of the NKG2D receptor complex and a second domain which functions
XX CC as a receptor or ligand. The polypeptide and its associated
XX CC polynucleotide are used for the preparation of a pharmaceutical
XX CC composition for the treatment of cancer, infections and/or autoimmune
XX CC conditions. The cancer may be a tumour of the head and neck, stomach,
XX CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,
XX CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,
XX CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.
XX CC The infectious diseases can be caused by viruses, bacteria, fungi,
XX CC protozoa or helminths. The autoimmune diseases include multiple
XX CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior
XX CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent
XX CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and
XX CC autoimmune hepatitis. The sequences represent DNA encoding the
XX CC polypeptides of the invention.
XX
XX SQ Sequence 771 BP; 181 A; 190 C; 206 G; 194 T; 0 other;
Query Match 24.6%; Score 107; DB 24; Length 771;
Best Local Similarity 99.0%; Pred. No. 5.5e-44;
Matches 207; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 163 TGGCTTCAGCAGGAACCCAGATGGAAGTATTTAAACGGCTGATCTACGCCACATCCAGTTTA 222
XX DB 517 TGGCTTCAGCAGGAACCCAGATGGAAGTATTTAAACGGCTGATCTACGCCACATCCAGTTTA 576
XX
XX QY 223 GGTTCCTGGTGTCCCAAAAGGTTTCAGTGGCAGTAGTCTGGGTCAGATTATTTCTTCACC 282
XX DB 577 GATTCCTGGTGTCCCAAAAGGTTTCAGTGGCAGTAGTCTGGGTCAGATTATTTCTTCACC 636
XX
XX QY 283 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAGCCCTATTACTGTCTACAATATGCTAGTTCT 342
XX DB 637 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAGACTATTACTGTCTACAATATGCTAGTTCT 696
XX
XX QY 343 CCGTACACGTTTCGGAGGGGGGACCAAGCT 371
XX DB 697 CCGTACACGTTTCGGAGGGGGGACCAAGCT 725
XX
XX RESULT 6
XX ID AAS97145
XX AC AAS97145 standard; DNA; 1497 BP.
XX
XX AC AAS97145;
XX

```

DT	26-FEB-2002	(first entry)
XX	3B10xP4-3	bispecific single chain Fv DNA.
XX	Human:	NGK2D; NKG2D receptor complex; cancer; infectious disease; tumour;
XX	KW	autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;
XX	KW	intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;
XX	KW	prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;
XX	KW	sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; ds;
XX	KW	helminth; cytostatic; antimicrobial; immunomodulatory; P4-2; P4-3; P4-14;
XX	KW	P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10xP4-3; 3B10xP4-14;
XX	KW	P53 tetramerisation domain; 3B10xP5-2; 3B10xP5-23.
OS	Homo sapiens.	
XX	WO200171005-A2.	
XX	27-SEP-2001.	
PD		
XX	26-MAR-2001;	2001WO-EP03414.
XX	24-MAR-2000;	2000EP-0106467.
XX	(KUFE/)	KUFE P.
PA	Kufer P,	Riethmueller G, Lutterbuese R, Borschert K, Kischel R;
PI	Mayer M,	Hofmeister R;
PI	WPI:	2002-055119/07.
XX	P-PSDB;	AAU72871.
DR		
DR		
XX		
PT	Multifunctional polypeptides comprising binding sites that specifically	
PT	recognise extracellular groups of the NKG2D receptor complex and	
PT	domains which function as receptors or ligands, useful for treating	
PT	cancers and infectious diseases	
XX		
PS	Example 5; Fig 16; 114pp; English.	
XX		
CC	The invention relates to a multifunctional polypeptide comprising a	
CC	domain with a binding site that specifically recognises an extracellular	
CC	group of the NKG2D receptor complex and a second domain which functions	
CC	as a receptor or ligand. The polypeptide and its associated	
CC	polynucleotide are used for the preparation of a pharmaceutical	
CC	composition for the treatment of cancer, infections and/or autoimmune	
CC	conditions. The cancer may be a tumour of the head and neck, stomach,	
CC	oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,	
CC	larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,	
CC	bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.	
CC	The infectious diseases can be caused by viruses, bacteria, fungi,	
CC	protozoa or helminths. The autoimmune diseases include multiple	
CC	sclerosis, grave's disease, ankylosing spondylitis, acute anterior	
CC	uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent	
CC	diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and	
CC	autoimmune hepatitis. The sequences represent DNA encoding the	
XX	polypeptides of the invention.	
XX		
SQ	Sequence 1497 BP; 364 A; 358 C; 407 G; 368 T; 0 other;	
	Query Match	24.68; Score 107; DB 24; Length 1497;
	Best Local Similarity	99.08; Pred. No. 5.4e-44;
	Matches 207; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	163	TGGCTTCAGCAGGAACCATGATGGAACCTATTAAACGCTGATCATCGCCACATCCAGTTTA 222
Db	1243	TGGCTTCAGCAGGAACCATGATGGAACCTATTAAACGCTGATCATCGCCACATCCAGTTTA 1302
QY	223	GGTTCCTGGTGTCCCCAAAGGTTTCAGTGGCAGTAGGTCTGGGTGCAGATTATTTCTTCACC 282
Db	1303	GATTCTGGTGTCCCCAAAGGTTTCAGTGGCAGTAGGTCTGGGTGCAGATTATTTCTTCACC 1362
QY	283	ATCAGCAGCCTTGAGTCTGAAGATTTTGTAGCCCTATTACTGTCTACAAATATGCTAGTTCT 342-
Db	1363	ATCAGCAGCCTTGAGTCTGAAGATTTTGTAGACTATTACTGTCTACAAATATGCTAGTTCT 1422

Qy	343	CCGTACACAGTTCGGAGGGGGGACCAAGCT	371
Db	1423	CCGTACACAGTTCGGAGGGGGGACCAAGCT	1451
RESULT 7			
AAAT77851			
ID	AAAT77851	standard; cDNA; 387 BP.	
XX			
AC	AAAT77851;		
XX			
DT	03-NOV-1997	(first entry)	
XX			
DE	Murine anti-human class II monoclonal antibody 44H104 VL chain cDNA		
XX			
KW	Antibody; light chain; variable region; hybridoma cell line 44H104;		
KW	immune response; enhance; stimulate; vaccine; immunodiagnosis;		
KW	antigen delivery; ss.		
XX			
OS	Mus musculus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..387	
FT		/tag= a	
FT		/note= "Encodes 44H104 light chain variable region,	
FT		including secretion signal; termination	
FT		codon not given"	
XX			
PN	WO9640941-A1.		
XX			
PD	19-DEC-1996.		
XX			
PF	07-JUN-1996;	96WO-CA00400.	
XX			
PR	07-JUN-1995;	95US-0483576.	
XX			
PA	(CONN-) CONNAUGHT LAB LTD.		
XX			
PI	Anand NN, Barber BH, Caterini JE, Cates GC, Klein MH;		
XX			
DR	WPI; 1997-077271/07.		
DR	P-PSDB; AAW22537.		
XX			
PT	Recombinant conjugate antibody mol., modified for delivering an		
PT	antigen - elicits enhanced immune response without the use of		
PT	adjuvant to generate antibodies which are useful in vaccines or		
XX	immuno:diagnosis		
XX			
PS	Example 1; Fig 1A; 64pp; English.		
XX			
CC	Novel recombinant conjugate antibody molecules comprise a monoclonal		
CC	antibody specific for a surface structure of antigen presenting		
CC	cells (APC), genetically modified to contain at least one antigen		
CC	exclusively at one or more preselected sites. The conjugate is capa		
CC	of delivering the antigen to APC and eliciting an immune response t		
CC	the antigen. The new conjugates are useful as vaccines and are able		
CC	to elicit an enhanced immune response without the use of an adjuvan		
CC	In a specific example, a conjugate was constructed using the murine		
CC	anti-human class II monoclonal antibody secreted by hybridoma		
CC	44H104. The peptide CLTB36 was chosen as antigen; it consists of		
CC	a tandemly linked T and B cell epitope derived from HIV MN strain,		
CC	The present sequence encodes the light chain variable region which		
CC	was PCR amplified from 44H104 and used in the preparation of a		
CC	conjugate with antigen CLTB36.		
XX			
SQ	Sequence 387 BP; 90 A; 95 C; 90 G; 112 T; 0 other;		
Query Match 23.4%; Score 102; DB 18; Length 387;			
Best Local Similarity 100.0%; Pred. No. 1.9e-41;			
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gap			
QV	42	GTITCCAGGTACGATGTGACATGCCAGATGACCGACGTCCTCATCTCTTATCTGCGC	

```

Db      48 GTTTCAGGTACCAGATGTGACATCCAGATGACCCAGTCTCCATCTCTTATCTGCCTC 107
      102 TCTGGGACAAAGAGTCAGTCTCACTTGTCTGGGCAAGTCAGGA 143
      108 TCTGGGACAAAGAGTCAGTCTCACTTGTCTGGGCAAGTCAGGA 149

RESULT 8
AAD32138
ID      AAD32138 standard; DNA; 387 BP.
XX
AC      AAD32138;
XX
DT      18-JUN-2002 (first entry)
XX
DE      Murine 44H104 mab variable light chain (VL) DNA.
XX
KW      Murine; mab; light chain; VL; conjugate antibody; antigen delivery;
KW      immune system; vaccine; detecting agent; antibacterial; gene; ds.
XX
OS      Mus sp.
XX
FH      Key
FH      CDS
FT      Location/Qualifiers
FT      1..387
FT      /product= "Murine 44H104 mab VL"
FT      /trans_except= (pos:277..279, aa:Thr)
FT      /note= "this translation exception occurs while
FT      decoding for murine mab VL alternative version
FT      (AAE20204); CDS does not include stop codon"
FT      /partial
XX
PN      US2002025315-A1.
XX
XX
PD      28-FEB-2002.
XX
PF      14-JAN-1998; 98US-0007093.
XX
PR      14-JAN-1998; 98US-0007093.
XX
PA      (ANAN/) ANAND N N.
PA      (BARB/) BARBER B H.
PA      (CATE/) CATES G A.
PA      (CATE/) CATERINI J E.
PA      (KLEI/) KLEIN M H.
XX
PI      Anand NN, Barber BH, Cates GA, Caterini JE, Klein MH;
XX
WPI; 2002-267519/31.
P-PSDB; AAE20200, AAE20204.
XX
Novel recombinant conjugate antibody, useful as a vaccine against
pathogens having a specific antigen, comprises a monoclonal antibody
specific for an antigen presenting cell surface structure -
XX
Example 1; Fig 1A; 28pp; English.
XX
The invention relates to a recombinant conjugate antibody, comprising
a monoclonal antibody specific for a surface structure of antigen
presenting cells genetically modified to contain an antigen moiety
for the purpose of delivery of the antigen moiety to antigen-
presenting cells of the immune system. The conjugate antibody is
formulated as a vaccine to protect a host against a disease caused
by a pathogen expressing the antigen. The antibody is useful as a
detecting agent. The present sequence is murine 44H104 mab variable
light chain (VL) DNA.
XX
SQ      Sequence 387 BP; 90 A; 95 C; 90 G; 112 T; 0 other;
Query Match      23.4%; Score 102; DB 24; Length 387;
Best Local Similarity 100.0%; Pred. No. 1.9e-41;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      42 GTTTCAGGTACCAGATGTGACATCCAGATGACCCAGTCTCCATCTCTTATCTGCCTC 101
      48 GTTTCAGGTACCAGATGTGACATCCAGATGACCCAGTCTCCATCTCTTATCTGCCTC 107
QY      102 TCTGGGACAAAGAGTCAGTCTCACTTGTCTGGGCAAGTCAGGA 143
      108 TCTGGGACAAAGAGTCAGTCTCACTTGTCTGGGCAAGTCAGGA 149
Db

RESULT 9
AAX00879
ID      AAX00879 standard; DNA; 276 BP.
XX
AC      AAX00879;
XX
DT      29-MAR-1999 (first entry)
XX
DE      Mouse derived RT3 phage antibody light chain pattern C genetic sequence.
XX
KW      Catalytic; antibody; phage display; immunising; phage expression vector;
KW      produg; scfv; ss.
XX
OS      Mus sp.
XX
FH      Key
FH      CDS
FT      Location/Qualifiers
FT      1..276
FT      /*tag= a
FT      /note= "the start and stop codons are not indicated"
XX
PN      US5855885-A.
XX
XX
PD      05-JAN-1999.
XX
PF      14-JUL-1994; 94US-0273146.
XX
PR      22-JAN-1993; 93US-0007684.
PR      14-JUL-1994; 94US-0273146.
XX
PA      (CHIS/) CHISWELL D.
PA      (DARS/) DARSLEY M J.
PA      (FITZ/) FITZGERALD K.
PA      (KENT/) KENTEN J H.
PA      (MART/) MARTIN M T.
PA      (MCCA/) MCCAFFERTY J.
PA      (SMIT/) SMITH R.
PA      (TITM/) TITMAS R C.
PA      (WILL/) WILLIAMS R O.
XX
PI      Chiswell D, Darsley MJ, Fitzgerald K, Kenten JH;
PI      Martin MT, McCafferty J, Smith R, Titmas RC, Williams RO;
XX
WPI; 1999-105036/09.
P-PSDB; AAW95480.
XX
Production of catalytic antibodies displayed on bacteriophages -
comprises generating a gene library of antibody-derived domains
inserting coding into a phage expression vector and isolating the
catalytic antibodies
XX
Example 4; Fig 11; 117pp; English.
XX
The invention relates to methods for producing catalytic antibodies
displayed on a phage. The method comprises: (a) generating a gene library
of antibody-derived domains; (b) inserting coding for the domains into a
phage expression vector; and (c) isolating the catalytic antibodies. The
phage expression vector incorporates a histidine peptide in tandem with a
myc peptide. The catalytic antibodies can be isolated by preparing an
antigen; optionally immunising an animal with the antigen; generating a
library of VH and VL domains from the immunised animal; cloning the VH
and VL domains into a phage expression vector to generate phage display
antibodies; selecting phage display antibodies which bind specifically
to the antigen; screening the selected phage display antibodies for

```

CC catalytic activity to substrate; and isolating the catalytic antibodies,  
 CC where the phage expression vector incorporates a histidine peptide in  
 CC tandem with a myc peptide. The processes are used to produce catalytic  
 CC antibodies, which can be used for in vivo activation of a prodrug. The  
 CC present sequence represents a genetic sequence of light chain PCR  
 CC pattern C from mouse derived RT3 phage antibodies.  
 XX  
 SQ Sequence 276 BP; 67 A; 65 C; 64 G; 80 T; 0 other;  
 Query Match 23.0%; Score 100; DB 20; Length 276;  
 Best Local Similarity 99.3%; Pred. No. 2e-40;  
 Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 163 TGGCTTCAGCAGGAGCAGATGGAACCTATTAAACGGCTGATCTACGCCACATCCAGTTTA 222  
 Db 79 TGGCTTCAGCAGGAGCAGATGGAACCTATTAAACGGCTGATCTACGCCACATCCAGTTTA 138  
 QY 223 GGTTCGGTGTCCCAAAAGGTTCACTGGCAGTAGTCTGGGTCAGATATTCTCTCACC 282  
 Db 139 GATTCGGTGTCCCAAAAGGTTCACTGGCAGTAGTCTGGGTCAGATATTCTCTCACC 198  
 QY 283 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAG 313  
 Db 199 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAG 229  
 RESULT 10  
 AAX00875  
 ID AAX00875 standard; DNA; 276 BP.  
 XX  
 AC AAX00875;  
 XX  
 DT 29-MAR-1999 (first entry)  
 XX  
 DE Mouse derived RT3 phage antibody light chain pattern A genetic sequence.  
 XX  
 KW Catalytic; antibody; phage display; immunising; phage expression vector;  
 KW prodrug; scfv; ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..276  
 FT /\*tag= a  
 FT /note= "the start and stop codons are not indicated"  
 XX  
 PN US5855885-A.  
 XX  
 PD 05-JAN-1999.  
 XX  
 PF 14-JUL-1994; 94US-0273146.  
 XX  
 PR 22-JAN-1993; 93US-0007684.  
 PR 14-JUL-1994; 94US-0273146.  
 XX  
 PA (CHIS/) CHISWELL D.  
 PA (DARS/) DARSLEY M J.  
 PA (FITZ/) FITZGERALD K.  
 PA (KENT/) KENTEN J H.  
 PA (MART/) MARTIN M T.  
 PA (MCCA/) MCCAFFERTY J.  
 PA (SMIT/) SMITH R.  
 PA (TITM/) TITMAS R C.  
 PA (WILL/) WILLIAMS R O.  
 XX  
 PI Chiswell D, Darsley MJ, Fitzgerald K, Kenten JH;  
 PI Martin WT, McCafferty J, Smith R, Titmas RC, Williams RO;  
 XX  
 DR WP1; 1999-105036/09.  
 DR P-PSDB; AAW95476.  
 XX  
 PT Production of catalytic antibodies displayed on bacteriophages -  
 PT comprises generating a gene library of antibody-derived domains

PT inserting coding into a phage expression vector and isolating the  
 XX catalytic antibodies  
 PS Example 4; Fig 9A-F; 117pp; English.  
 XX  
 CC The invention relates to methods for producing catalytic antibodies  
 CC displayed on a phage. The method comprises: (a) generating a gene library  
 CC of antibody-derived domains; (b) inserting coding for the domains into a  
 CC phage expression vector; and (c) isolating the catalytic antibodies. The  
 CC phage expression vector incorporates a histidine peptide in tandem with a  
 CC myc peptide. The catalytic antibodies can be isolated by preparing an  
 CC antigen; optionally immunising an animal with the antigen; generating a  
 CC library of VH and VL domains from the immunised animal; cloning the VH  
 CC and VL domains into a phage expression vector to generate phage display  
 CC antibodies; selecting phage display antibodies which bind specifically  
 CC to the antigen; screening the selected phage display antibodies for  
 CC catalytic activity to substrate; and isolating the catalytic antibodies,  
 CC where the phage expression vector incorporates a histidine peptide in  
 CC tandem with a myc peptide. The processes are used to produce catalytic  
 CC antibodies, which can be used for in vivo activation of a prodrug. The  
 CC present sequence represents a genetic sequence of light chain pattern A  
 CC from mouse derived RT3 phage antibodies.  
 XX  
 SQ Sequence 276 BP; 68 A; 65 C; 63 G; 80 T; 0 other;  
 Query Match 23.0%; Score 100; DB 20; Length 276;  
 Best Local Similarity 99.3%; Pred. No. 2e-40;  
 Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 163 TGGCTTCAGCAGGAGCAGATGGAACCTATTAAACGGCTGATCTACGCCACATCCAGTTTA 222  
 Db 79 TGGCTTCAGCAGGAGCAGATGGAACCTATTAAACGGCTGATCTACGCCACATCCAGTTTA 138  
 QY 223 GGTTCGGTGTCCCAAAAGGTTCACTGGCAGTAGTCTGGGTCAGATATTCTCTCACC 282  
 Db 139 GATTCGGTGTCCCAAAAGGTTCACTGGCAGTAGTCTGGGTCAGATATTCTCTCACC 198  
 QY 283 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAG 313  
 Db 199 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAG 229  
 RESULT 11  
 AAN30165  
 ID AAN30165 standard; DNA; 450 BP.  
 XX  
 AC AAN30165;  
 XX  
 DT 25-MAY-1992 (first entry)  
 XX  
 DE Sequence encoding the leader, variable region and first 16 AAs of  
 DE the constant region of the kappa-chain (light chain) of MOPC41.  
 XX  
 KW Diagnosis; therapy; immunoglobulin; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 13..78  
 FT /\*tag= a  
 FT /product= leader  
 FT CDS 79..402  
 FT /\*tag= b  
 FT /product= variable region  
 FT CDS 403..450  
 FT /\*tag= c  
 FT /product= constant region  
 XX  
 PN EP88994-A.  
 XX  
 PD 21-SEP-1983.  
 XX  
 PF 10-MAR-1983; 83EP-0001655.

```

XX 15-MAR-1982; 82US-0358414.
PR 05-DEC-1983; 83US-0558551.
XX
XX (SCHE ) SCHERING CORP.
PA (DNAX-) DNAX RES INST.
XX
XX Moore KW, Zaffaroni A;
PI
XX WPI; 1983-772290/39.
DR P-PSDB; AAP30251.
XX
XX Transformed expression vectors or plasmid(s) - with double
PT stranded DNA sequence coding only for desired part of polypeptide
PT chain
XX
XX Example; Page 40-41; 68pp; English.
XX
XX The pref. vector or plasmid of the invention has a double-stranded
CC DNA seq. coding for a variable region of a light or heavy chain of
CC IgG, or for a variable region of a light or heavy chain of an
CC immunoglobulin specific for an enzyme or surface protein. The
CC sequence esp. codes for a variable region of a light chain having 95-
CC 115 AAs or for a variable region of a heavy chain having 110-125 AAs
CC esp. including the D region of the heavy chain.
XX
XX Sequence 450 BP; 108 A; 114 C; 104 G; 124 T; 0 other;
SQ
Query Match 23.0%; Score 100; DB 4; Length 450;
Best Local Similarity 99.3%; Pred. No. 1.9e-40;
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 163 TGGCTTCAGCAGGAACACAGATGAAGTATTAACGGCTGATCTACGCCACATCCAGTTTA 222
DB 181 TGGCTTCAGCAGGAACACAGATGAAGTATTAACGGCTGATCTACGCCACATCCAGTTTA 240
QY 223 GGTCTGGTGTCGCCAAAGGTTTCAGTGGCAGTAGGTCTGGGTGAGATTATTCCTCACC 282
DB 241 GATTCTGGTGTCGCCAAAGGTTTCAGTGGCAGTAGGTCTGGGTGAGATTATTCCTCACC 300
QY 283 ATCAGCAGCCTTCAGTCTGAAGATTTTGTAG 313
DB 301 ATCAGCAGCCTTCAGTCTGAAGATTTTGTAG 331
RESULT 12
AAV20086
ID AAV20086 standard; DNA; 535 BP.
XX
XX AAV20086;
XX
XX 14-JUL-1998 (first entry)
XX
XX Consensus DNA sequence of the murine variable light chain region.
XX
XX Mouse; Act-1 antibody; human alpha4-beta7 integrin;
XX Muscosal adressin cell adhesion molecule-1; MadCAM-1;
XX humanised antibody; murine antigen binding region; inhibition;
XX leukocyte infiltration of tissue; treatment; inflammatory disease;
XX inflammatory bowel disease; ss.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX CDS 16..435
XX /*tag- a
XX /note- "no stop codon given"
XX sig_peptide 16..75
XX /*tag- b
XX mat_peptide 76..435
XX /*tag- c
XX
XX WO9806248-A2.

```

```

XX 19-FEB-1998.
XX
XX 06-AUG-1997; 97WO-US13884.
XX
XX 15-AUG-1996; 96US-0700737.
XX
XX (LEUK-) LEUKOSITE INC.
XX
XX Bendig MM, Jones ST, Newman W, Ponath PD, Ringler DJ;
XX Saldanha J;
XX
XX WPI; 1998-159172/14.
DR P-PSDB; AAW53817.
XX
XX Humanised immunoglobulin reactive with alpha-4-beta-7 integrin -
PT used for treating inflammatory disease, pancreatitis, diabetes,
PT asthma, graft versus host disease and sarcoidosis
XX
XX Example 1; Fig 3; 145pp; English.
XX
XX The present sequence represents the consensus nucleotide sequence
CC comprising the variable region of murine Act-1 antibody determined from
CC several independent mouse light chain variable region clones. Act-1 is
CC active against human alpha4-beta7 integrin. Muscosal adressin cell
CC adhesion molecule-1 (MadCAM-1) is a ligand of this particular integrin.
CC The Act-1 antibody interferes with alpha4-beta7 integrin binding to
CC MadCAM-1, which is present of high endothelial venules in muscosal
CC lymph nodes. Variable regions were amplified from DNA encoding Act-1
CC using degenerate PCR primers AAV20083-84. The degeneracy of the PCR
CC primers produced several different sequences, of which the present
CC sequence is a consensus sequence. The present sequence was used to
CC construct chimeric, humanised Act-1 antibodies, which contain murine
CC antigen binding regions. The humanised immunoglobulin can be used to
CC inhibit the interaction of cells bearing alpha4-beta7 with cells bearing
CC a ligand for alpha4-beta7. It can be used for inhibiting leukocyte
CC infiltration of tissues, e.g. for treating inflammatory diseases such
CC as inflammatory bowel disease. The immunoglobulin can also be used for
CC detection, isolation and diagnosis.
XX
XX Sequence 535 BP; 126 A; 128 C; 132 G; 149 T; 0 other;
SQ
Query Match 21.4%; Score 93; DB 19; Length 535;
Best Local Similarity 100.0%; Pred. No. 6.7e-37;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 343 CCGTACACGTTCCGAGGGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACT 402
DB 373 CCGTACACGTTCCGAGGGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACT 432
QY 403 GTATCCATCTTCCACCATCCAGTAAGCTTTGGG 435
DB 433 GTATCCATCTTCCACCATCCAGTAAGCTTTGGG 465
RESULT 13
AAQ90431
ID AAQ90431 standard; DNA; 438 BP.
XX
XX AAQ90431;
XX
XX 02-FEB-1996 (first entry)
XX
XX DNA encoding anti-idiotypic antibody Idiol7 clone 17Rb1.
XX
XX Antibody; cancer; CDR; heavy chain; light chain; immunoglobulin;
XX complementarity determining region, ds.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX CDS 1..438
XX /*tag- a

```

FT sig\_peptide /product= anti-idiotypic\_antibody\_idio17  
FT 1..39  
FT /\*tag= b

XX JP07101999-A.

PN 18-APR-1995.

XX 06-OCT-1993; 93JP-0272950.

XX 06-OCT-1993; 93JP-0272950.

XX (HAGI/) HAGIWARA Y.

XX WPI; 1995-182987/24.

XX P-PSDB; AAR74956.

XX Novel anti-idiotypic antibody against an human anticancer monoclonal  
FT antibody - and DNA sequences encoding the antibody, useful in  
FT pharmacology, medicine and biochemical fields.

XX Example 5; Page 19; 28pp; Japanese.

XX AAQ90425-Q90434 are DNA molecules encoding anti-idiotypic antibodies  
CC idio3, idio17, idio20, idio27 and idio33 against a human anticancer  
CC monoclonal antibody. These antibodies and DNA encoding them are useful  
CC in pharmacological, medical and biochemical fields of research.

XX Sequence 438 BP; 104 A; 120 C; 112 G; 102 T; 0 other;

Query Match 20.7%; Score 90; DB 16; Length 438;  
Best Local Similarity 100.0%; Pred. No. 2.2e-35;  
Matches. 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 346 TACACGTCGGAGGGGACCAAGCTGGAAATAAAGCGGCTGATGCTGCACCACTGTA 405

Db 334 TACACGTCGGAGGGGACCAAGCTGGAAATAAAGCGGCTGATGCTGCACCACTGTA 393

QY 406 TCCATCTCCACCATCCAGTAAGCTTGGG 435

Db 394 TCCATCTCCACCATCCAGTAAGCTTGGG 423

RESULT 14

ABS52791

ID ABS52791 standard; cDNA; 979 BP.

XX AC ABS52791;

XX 29-NOV-2002 (first entry)

XX Murine m166 antibody light chain cDNA.

DE Mouse; m166 light chain; gene; ss; antibacterial; PcrV antigen;  
KW Pseudomonas aeruginosa infection.

XX Mus sp.

XX KEY Location/Qualifiers

FT CDS 44..748

FT /\*tag= a

FT /product= "Murine m166 light chain"

XX WO200264161-A2.

XX 22-AUG-2002.

XX 25-JAN-2002; 2002WO-US02382.

XX 26-JAN-2001; 2001US-0770916.

XX 29-JAN-2001; 2001US-264795P.

XX (MCWR-) MCW RES FOUND INC.

PA

PA (REGC ) UNIV CALIFORNIA.

XX Frank DW, Wiener-Kronish J, Yahr TL, Sawa T, Fritz RB;

XX WPI; 2002-636608/68.

DR P-PSDB; ABG31731.

XX Inhibiting Pseudomonas aeruginosa infection by inoculating a patient  
FT with PcrV antigen, or inoculating a patient with a gene vaccine  
FT encoding the PcrV antigen

XX Example 6; Fig 7; 63pp; English.

XX The invention relates to a method of inhibiting Pseudomonas aeruginosa  
CC infection by inoculating a patient with PcrV antigen, or inoculating a  
CC patient with a gene vaccine encoding the PcrV antigen. The method is  
CC useful for inhibiting P.aeruginosa infection in a patient, preferably a  
CC human patient, for treating P.aeruginosa infection, where the infected  
CC patient is inoculated with PcrV antigen, and for diagnosing P.aeruginosa  
CC infection. An antibody to the antigen, or its fragments, (preferably  
CC monoclonal antibodies) is useful for modulating cytotoxicity of  
CC Pseudomonas to a human cell which involves contacting the Pseudomonas  
CC with the antibody in the presence of the human cell. A pharmaceutical  
CC composition comprising the monoclonal antibodies can be used for treating  
CC or preventing Pseudomonas infection, or for reducing pathogenicity of  
CC Pseudomonas in a patient. This sequence represents cDNA encoding a murine  
CC m166 antibody light chain polypeptide of the invention.

XX Sequence 979 BP; 285 A; 259 C; 204 G; 231 T; 0 other;

Query Match 19.8%; Score 86; DB 24; Length 979;  
Best Local Similarity 100.0%; Pred. No. 2.3e-33;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 CTCGCTACAGCTTCGGAGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAA 400

Db 384 CTCGCTACAGCTTCGGAGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAA 443

QY 401 CTGTATCCCATCTTCCCACCATCCAGT 426

Db 444 CTGTATCCCATCTTCCCACCATCCAGT 469

RESULT 15

AAT85091

ID AAT85091 standard; cDNA; 642 BP.

XX AC AAT85091;

XX 18-NOV-1997 (first entry)

XX Mouse monoclonal antibody B9 light chain encoding cDNA.

DE Human plasma apolipoprotein B-100; arteriosclerotic lipoprotein;  
KW antibody; Fab; ds.

XX Mus musculus.

XX KEY Location/Qualifiers

FT CDS 1..642

FT /\*tag= a

FT /product= B9L

FT /note= "Stop codon not shown"

FT misc\_feature 322..642

FT /\*tag= b

FT /label= Ckappa

XX JP09154587-A.

XX 17-JUN-1997.

XX 09-MAY-1996; 96JP-0114492.

XX

PR 02-NOV-1995; 95KR-0039459.  
XX (KOAD ) KOREAN SCI & TECHNOLOGY RES CENT.  
XX WPI; 1997-367067/34.  
XX P-PSDB; AAW27089.  
XX DNA encoding mouse antibody binding human plasma apo-lipoprotein  
PT B-100 - useful for removing arteriosclerotic lipoprotein(s)  
XX Claim 4; Fig 6; 17pp; Japanese.  
XX The present sequence encodes the mouse monoclonal antibody B9 light  
CC chain (B9L), which binds specifically to human blood apolipoprotein  
CC B-100. The nucleic acid can be used in a method for the preparation  
CC of a reconstituted antibody which specifically binds human plasma  
CC apolipoprotein B-100. The antibody can be used as a reagent for  
CC determining the concentration of human plasma apolipoprotein B-100  
CC in a sample. The antibody is also useful in a drug composition for  
CC selectively removing arteriosclerotic lipoproteins containing human  
XX plasma apolipoprotein B-100.  
SQ Sequence 642 BP; 188 A; 152 C; 152 G; 150 T; 0 other;  
Query Match 19.5%; Score 85; DB 18; Length 642;  
Best Local Similarity 100.0%; Pred. No. 7.4e-33;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 342 TCCGTACACGTTCCGGAGGGGGACCAAGCTGGAAATAAAGCGGCTGATGCTGCACCAAC 401  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
282 TCCGTACACGTTCCGGAGGGGGACCAAGCTGGAAATAAAGCGGCTGATGCTGCACCAAC 341  
QY 402 TGTATCCATCTTCCACCATCCAGT 426  
Db ||||||||||||||||||||||||||||  
342 TGTATCCATCTTCCACCATCCAGT 366  
Search completed: August 30, 2003, 19:57:02  
Job time : 191.312 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 19:44:55 ; Search time 158.756 Seconds  
(without alignments)  
6286.388 Million cell updates/sec

Title: US-08-836-455-1  
Perfect score: 435  
Sequence: 1 ATGGGGGCCCTGCTCAGAT.....CACCATCCAGTAAGCTTGGG 435

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1533700 seqs, 1147125425 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications\_NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	435	100.0	435	9	US-09-861-294-1
2	435	100.0	435	12	US-10-367-506-1
3	108	24.8	402	8	US-08-779-784-5
4	102	23.4	387	9	US-09-007-093-1
5	100	23.0	351	8	US-08-779-784-17
6	85	19.5	698	13	US-10-006-773-18
7	82	18.9	324	10	US-09-924-099-11
8	82	18.9	407	10	US-09-924-099-27
9	82	18.9	711	10	US-09-924-099-19
10	82	18.9	729	10	US-09-924-099-20
11	80	18.4	729	13	US-10-006-773-10
12	79	18.2	351	8	US-08-779-784-16
13	79	18.2	426	10	US-09-840-459-95
14	79	18.2	426	10	US-09-840-459-101
15	77	17.7	504	13	US-10-006-773-14
16	73	16.8	390	12	US-10-268-883-10

17	14.7	363	11	US-09-929-665-9	Sequence 9, Appli
18	14.7	363	11	US-09-929-665-10	Sequence 10, Appli
19	14.7	363	11	US-09-929-546-9	Sequence 9, Appli
20	14.7	363	11	US-09-929-546-10	Sequence 10, Appli
21	14.0	739	10	US-09-887-803-5	Sequence 5, Appli
22	13.1	390	13	US-10-146-305-5	Sequence 5, Appli
23	13.1	447	10	US-09-990-205-1	Sequence 1, Appli
24	13.1	447	14	US-10-153-401-1	Sequence 1, Appli
25	12.4	443	11	US-09-967-719C-3	Sequence 3, Appli
26	11.0	51	8	US-08-779-784-18	Sequence 18, Appli
27	11.0	391	11	US-09-726-258-34	Sequence 34, Appli
28	11.0	705	12	US-10-281-479A-22	Sequence 22, Appli
29	11.0	714	14	US-10-216-484-10	Sequence 10, Appli
30	11.0	831	10	US-09-903-327A-3	Sequence 3, Appli
31	11.0	3344	14	US-10-176-380-17	Sequence 17, Appli
32	11.0	5691	10	US-09-897-006-11	Sequence 11, Appli
33	11.0	5691	11	US-09-897-511A-11	Sequence 11, Appli
34	11.0	5711	10	US-09-897-006-8	Sequence 8, Appli
35	11.0	5711	11	US-09-897-511A-8	Sequence 8, Appli
36	11.0	6255	10	US-09-897-006-13	Sequence 13, Appli
37	11.0	6255	11	US-09-897-511A-13	Sequence 13, Appli
38	10.3	384	9	US-09-905-243-20	Sequence 20, Appli
39	10.3	737	10	US-09-919-344-7	Sequence 7, Appli
40	10.1	364	10	US-09-797-941A-3	Sequence 3, Appli
41	10.1	387	9	US-09-905-243-27	Sequence 27, Appli
42	10.1	472	9	US-09-797-481-5	Sequence 5, Appli
43	9.9	358	10	US-09-800-908-10	Sequence 10, Appli
44	9.7	351	11	US-09-469-485-3	Sequence 3, Appli
45	9.7	393	14	US-10-195-752-2	GENERAL INFORMA

ALIGNMENTS

RESULT 1

US-09-861-294-1  
; Sequence 1, Application US/09861294  
; Patent No. US20020098190A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; FILE OF INVENTION: TUMORS BEARING HMF AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 435  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(435)  
; NAME/KEY: sig\_peptide  
; LOCATION: (1)...(60)  
; NAME/KEY: mat\_peptide  
; LOCATION: (61)...(435)  
US-09-861-294-1

Query Match 100.0%; Score 435; DB 9; Length 435;  
Best Local Similarity 100.0%; Pred. No. 4.6e-232;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGGGGGCCCCGCTCAGATTCTTGGTCTCTGTTTCCAGGTACCATGT 60  
Db 1 ATGGGGGGCCCCGCTCAGATTCTTGGTCTCTGTTTCCAGGTACCATGT 60

QY 61 GACATCCAGATGACCCAGTCTCCATCCTCCTTATCTGCTCTGCGACAAAGAGTCAGT 120  
DB 61 GACATCCAGATGACCCAGTCTCCATCCTCCTTATCTGCTCTGCGACAAAGAGTCAGT 120  
QY 121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180  
DB 121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180  
QY 181 GATGGAACCTATTAAACGCTGATCTACGCCACATCCAGTTTCTGTTGTCGCCAAA 240  
DB 181 GATGGAACCTATTAAACGCTGATCTACGCCACATCCAGTTTCTGTTGTCGCCAAA 240  
QY 241 AGTTTCAGTGGCAGTAGTCTGGGTGAGATTATTTCTTCACCATCAGCAGCCTTGAAGTCT 300  
DB 241 AGTTTCAGTGGCAGTAGTCTGGGTGAGATTATTTCTTCACCATCAGCAGCCTTGAAGTCT 300  
QY 301 GAAGATTTTGTAGCTTATTACTGTCTACAATATGCTTCTCGGTACACGTTCCGGAGGG 360  
DB 301 GAAGATTTTGTAGCTTATTACTGTCTACAATATGCTTCTCGGTACACGTTCCGGAGGG 360  
QY 361 GGGACCAAGCTGGAATAAAACGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
DB 361 GGGACCAAGCTGGAATAAAACGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
QY 421 TCCAGTAAGCTTGGG 435  
DB 421 TCCAGTAAGCTTGGG 435

## RESULT 2

US-10-367-506-1  
; Sequence 1, Application US/10367506  
; Publication No. US20030152575A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TITLE OF INVENTION: TUMORS BEARING HMGF AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/10/367,506  
; CURRENT FILING DATE: 2003-02-13  
; PRIOR APPLICATION NUMBER: US/09/861,294  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 435  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(435)  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: (1)...(60)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (61)...(435)  
US-10-367-506-1

Query Match 100.0%; Score 435; DB 12; Length 435;  
Best Local Similarity 100.0%; Pred. No. 4.6e-232;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGCCCCGTCTCAGATTCCTTGGTCTTGTGCTTGTTCAGGTACAGATGT 60  
DB 1 ATGGGGGCCCCGTCTCAGATTCCTTGGTCTTGTGCTTGTTCAGGTACAGATGT 60  
QY 61 GACATCCAGATGACCCAGTCTCCATCCTCCTTATCTGCTCTGCGACAAAGAGTCAGT 120

DB 61 GACATCCAGATGACCCAGTCTCCATCCTCCTTATCTGCTCTGCGACAAAGAGTCAGT 120  
QY 121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180  
DB 121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180  
QY 181 GATGGAACCTATTAAACGCTGATCTACGCCACATCCAGTTTCTGTTGTCGCCAAA 240  
DB 181 GATGGAACCTATTAAACGCTGATCTACGCCACATCCAGTTTCTGTTGTCGCCAAA 240  
QY 241 AGTTTCAGTGGCAGTAGTCTGGGTGAGATTATTTCTTCACCATCAGCAGCCTTGAAGTCT 300  
DB 241 AGTTTCAGTGGCAGTAGTCTGGGTGAGATTATTTCTTCACCATCAGCAGCCTTGAAGTCT 300  
QY 301 GAAGATTTTGTAGCTTATTACTGTCTACAATATGCTTCTCGGTACACGTTCCGGAGGG 360  
DB 301 GAAGATTTTGTAGCTTATTACTGTCTACAATATGCTTCTCGGTACACGTTCCGGAGGG 360  
QY 361 GGGACCAAGCTGGAATAAAACGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
DB 361 GGGACCAAGCTGGAATAAAACGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
QY 421 TCCAGTAAGCTTGGG 435  
DB 421 TCCAGTAAGCTTGGG 435

## RESULT 3

US-08-779-784-5  
; Sequence 5, Application US/08779784  
; Publication No. US20020164325A1  
; GENERAL INFORMATION:  
; APPLICANT: Rodriguez, Moses  
; APPLICANT: Miller, David J.  
; APPLICANT: Asakura, Kunihiko  
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM  
; TITLE OF INVENTION: REMYELINATION USING MONOCLONAL AUTOANTIBODIES  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; STREET: Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/779,784  
; FILING DATE: 07-JAN-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/692,084  
; FILING DATE: 08-AUG-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/236,520  
; FILING DATE: 29-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1694  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 402 base pairs  
; TYPE: nucleic acid

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; US-08-779-784-5

Query Match      24.8%; Score 108; DB 8; Length 402;
Best Local Similarity 99.4%; Pred. No. 9.2e-50;
Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 235 CCACAAAGGTCAGTGGCAGTAGGCTGGGTGACAGATTATCTCTACCATCAGCAGCCTT 294
Db 241 CCCAAAGGTTAGTGGCAGTAGGCTGGGTGACAGATTATCTCTACCATCAGCAGCCTT 300
QY 295 GAGTCTGAAGATTCTTCTAGCCTATTACTGTCTACAAATATGCTAGTCTCGGTACAGTTC 354
Db 301 GAGTCTGAAGATTCTTCTAGCCTATTACTGTCTACAAATATGCTAGTCTCGGTACAGTTC 360
QY 355 GGAGGGGGGACCAAGCTGGAATAAATAACGGGCTGATGCT 393
Db 361 GGAGGGGGGACCAAGCTGGAATAAATAACGGGCTGATGCT 399

RESULT 4
US-09-007-093-1
; Sequence 1, Application US/09007093
; Patent No. US2002025315A1
; GENERAL INFORMATION:
; APPLICANT: Anand, Naveen N
; APPLICANT: Barber, Brian H
; APPLICANT: Cates, George A
; APPLICANT: Caterini, Judith E
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF
; TITLE OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/007,093
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,576
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-765
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-007-093-1

Query Match      23.4%; Score 102; DB 9; Length 387;
```

```
Best Local Similarity 100.0%; Pred. No. 2e-46;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 GTTTCAGGTACCAGATGTGACATCAGATGACCCAGTCTCCATCTCTTATCTGCCTC 101
Db 48 GTTTCAGGTACCAGATGTGACATCAGATGACCCAGTCTCCATCTCTTATCTGCCTC 107
QY 102 TCTGGGACAAAGAGTCAGTCTCACTTGTCTGGGCAAGTCAGGA 143
Db 108 TCTGGGACAAAGAGTCAGTCTCACTTGTCTGGGCAAGTCAGGA 149

RESULT 5
US-08-779-784-17
; Sequence 17, Application US/08779784
; Publication No. US20020164325A1
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Moses
; APPLICANT: Miller, David J.
; APPLICANT: Asakura, Kunihiro
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM
; TITLE OF INVENTION: REMYELINATION USING MONOCLONAL AUTOANTIBODIES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,784
; FILING DATE: 07-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/692,084
; FILING DATE: 08-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/236,520
; FILING DATE: 29-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-08-779-784-17

Query Match      23.0%; Score 100; DB 8; Length 351;
Best Local Similarity 99.3%; Pred. No. 2.6e-45;
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 163 TGGCTTCAGCAGGACCAAGATGGAAGTATTAACGCTGATCTACGCCATCCAGTTTA 222
Db 169 TGGCTTCAGCAGGACCAAGATGGAAGTATTAACGCTGATCTACGCCATCCAGTTTA 228
QY 223 GGTTCTGGTGTCCCAAAAGGTTTCACTGGCAGTAGTCTGGGTACAGATTATCTCTCACC 282
```

Db 229 GATTCTGGTGTCCCAAAAGGTTAGTGGCAGTAGGCTGGTCCAGATTATTCTCTCACC 288  
Qy 283 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAG 313  
Db 289 ATCAGCAGCCTTGAGTCTGAAGATTTTGTAG 319

RESULT 6  
US-10-006-773-18  
; Sequence 18, Application US/10006773  
; Publication No. US20020132983A1  
; GENERAL INFORMATION:  
; APPLICANT: Junghans, Richard P.  
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Ant  
; FILE REFERENCE: 003  
; CURRENT APPLICATION NUMBER: US/10/006,773  
; PRIOR FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: 60/250,089  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 698  
; TYPE: DNA  
; ORGANISM: Mus sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (6)..(401)  
; OTHER INFORMATION: 3e11 Light chain V region, plus leader  
US-10-006-773-18

Query Match 19.5%; Score 85; DB 13; Length 698;  
Best Local Similarity 100.0%; Pred. No. 6.2e-37;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 342 TCCGTACACGTCGGAGGGGACCAAGCTGGAATAAAACGGCTGATGCTGCACCAAC 401  
Db 359 TCCGTACACGTCGGAGGGGACCAAGCTGGAATAAAACGGCTGATGCTGCACCAAC 418

Qy 402 TGTATCCATCTTCCACCATCCAGT 426  
Db 419 TGTATCCATCTTCCACCATCCAGT 443

RESULT 7  
US-09-924-099-11  
; Sequence 11, Application US/09924099  
; Patent No. US20020128450A1  
; GENERAL INFORMATION:  
; APPLICANT: NISHIDA, Yoshihiro  
; APPLICANT: OKURA, Takanori  
; APPLICANT: TANIMOTO, Tadao  
; APPLICANT: KURIMOTO, Masashi  
; TITLE OF INVENTION: PEPTIDE  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/924,099  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 33  
; SEQ ID NO 11  
; LENGTH: 324  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-924-099-11

Query Match 18.9%; Score 82; DB 10; Length 324;  
Best Local Similarity 99.2%; Pred. No. 2.9e-35;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 241 AGGTTCAAGTGGCAGTAGGCTGGGTCTCAGATTATTCTCTCACCATCAGCAGCCTTGAGTCT 300  
Db 181 AGGTTCAAGTGGCAGTAGGCTGGGTCTCAGATTATTCTCTCACCATCAGCAGCCTTGAGTCT 240

Qy 301 GAAGATTTTCTAGCCTATTACTGTCTACAATATGCTAGTCTCCGTACACGTTCCGGAGGG 360  
Db 241 GAAGATTTTCTAGCCTATTACTGTCTACAATATGCTAGTCTCCGTACACGTTCCGGAGGG 300

Qy 361 GGGACCAAGCTGG 373  
Db 301 GGGACCAAGCTGG 313

RESULT 8  
US-09-924-099-27  
; Sequence 27, Application US/09924099  
; Patent No. US20020128450A1  
; GENERAL INFORMATION:  
; APPLICANT: NISHIDA, Yoshihiro  
; APPLICANT: OKURA, Takanori  
; APPLICANT: TANIMOTO, Tadao  
; APPLICANT: KURIMOTO, Masashi  
; TITLE OF INVENTION: PEPTIDE  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/924,099  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 33  
; SEQ ID NO 27  
; LENGTH: 407  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(407)  
; NAME/KEY: sig peptide  
; LOCATION: (1)..(60)  
US-09-924-099-27

Query Match 18.9%; Score 82; DB 10; Length 407;  
Best Local Similarity 99.2%; Pred. No. 2.9e-35;  
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 241 AGGTTCAAGTGGCAGTAGGCTGGGTCTCAGATTATTCTCTCACCATCAGCAGCCTTGAGTCT 300  
Db 241 AGGTTCAAGTGGCAGTAGGCTGGGTCTCAGATTATTCTCTCACCATCAGCAGCCTTGAGTCT 300

Qy 301 GAAGATTTTCTAGCCTATTACTGTCTACAATATGCTAGTCTCCGTACACGTTCCGGAGGG 360  
Db 301 GAAGATTTTCTAGCCTATTACTGTCTACAATATGCTAGTCTCCGTACACGTTCCGGAGGG 360

Qy 361 GGGACCAAGCTGG 373  
Db 361 GGGACCAAGCTGG 373

RESULT 9  
US-09-924-099-19  
; Sequence 19, Application US/09924099  
; Patent No. US20020128450A1  
; GENERAL INFORMATION:  
; APPLICANT: NISHIDA, Yoshihiro  
; APPLICANT: OKURA, Takanori  
; APPLICANT: TANIMOTO, Tadao  
; APPLICANT: KURIMOTO, Masashi  
; TITLE OF INVENTION: PEPTIDE  
; FILE REFERENCE:

```
; CURRENT APPLICATION NUMBER: US/09/924,099
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 19
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial DNA to code for the amino acid sequence of
; OTHER INFORMATION: SEQ
; OTHER INFORMATION: ID NO:9
US-09-924-099-19

Query Match      18.9%; Score 82; DB 10; Length 711;
Best Local Similarity 99.2%; Pred. No. 2.9e-35;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 241 AGGTTTCAGTGGCAGTAGGTCGGTCAGATTTCTTCACCATCAGCAGCCTTGAGTCT 300
DB 571 AGGTTTCAGTGGCAGTAGGTCGGTCAGATTTCTTCACCATCAGCAGCCTTGAGTCT 630
QY 301 GAAGATTTTGTAGCTATTACTGTCTACAATATGCTTCCGTCACAGTTCGGAGGG 360
DB 631 GAAGATTTTGTAGACTATTACTGTCTACAATATGCTTCCGTCACAGTTCGGAGGG 690
QY 361 GGGACCAAGCTGG 373
DB 691 GGGACCAAGCTGG 703

RESULT 10
US-09-924-099-20
; Sequence 20, Application US/09924099
; Patent No. US20020128450A1
; GENERAL INFORMATION:
; APPLICANT: NISHIDA, Yoshihiro
; APPLICANT: OKURA, Takanori
; APPLICANT: TANIMOTO, Tadao
; APPLICANT: KURIMOTO, Masashi
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/924,099
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 20
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial DNA to code for the amino acid sequence of
; OTHER INFORMATION: SEQ
; OTHER INFORMATION: ID NO:10
US-09-924-099-20

Query Match      18.9%; Score 82; DB 10; Length 729;
Best Local Similarity 99.2%; Pred. No. 2.9e-35;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 241 AGGTTTCAGTGGCAGTAGGTCGGTCAGATTTCTTCACCATCAGCAGCCTTGAGTCT 300
DB 571 AGGTTTCAGTGGCAGTAGGTCGGTCAGATTTCTTCACCATCAGCAGCCTTGAGTCT 630
QY 301 GAAGATTTTGTAGCTATTACTGTCTACAATATGCTTCCGTCACAGTTCGGAGGG 360
DB 631 GAAGATTTTGTAGACTATTACTGTCTACAATATGCTTCCGTCACAGTTCGGAGGG 690
QY 361 GGGACCAAGCTGG 373
DB 691 GGGACCAAGCTGG 703

RESULT 11
US-10-006-773-10
; Sequence 10, Application US/10006773
; Publication No. US20020132983A1
; GENERAL INFORMATION:
; APPLICANT: Junghans, Richard P.
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor A
; FILE REFERENCE: 003
; CURRENT APPLICATION NUMBER: US/10/006,773
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: 60/250,089
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(410)
; OTHER INFORMATION: 3D8 Light chain V region, plus leader
US-10-006-773-10

Query Match      18.4%; Score 80; DB 13; Length 729;
Best Local Similarity 100.0%; Pred. No. 3.8e-34;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 347 ACACGTTTCGGAGGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGTAT 406
DB 373 ACACGTTTCGGAGGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGTAT 432
QY 407 CCATCTTCCACCATCCAGT 426
DB 433 CCATCTTCCACCATCCAGT 452

RESULT 12
US-08-779-784-16
; Sequence 16, Application US/08779784
; Publication No. US20020164325A1
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Moses
; APPLICANT: Miller, David J.
; APPLICANT: Asakura, Kunihiko
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,784
```

```
QY 301 GAAGATTTTGTAGCCTATTACTGTCTACAATATGCTTCCGTCACAGTTCGGAGGG 360
DB 631 GAAGATTTTGTAGACTATTACTGTCTACAATATGCTTCCGTCACAGTTCGGAGGG 690
QY 361 GGGACCAAGCTGG 373
DB 691 GGGACCAAGCTGG 703

RESULT 11
US-10-006-773-10
; Sequence 10, Application US/10006773
; Publication No. US20020132983A1
; GENERAL INFORMATION:
; APPLICANT: Junghans, Richard P.
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor A
; FILE REFERENCE: 003
; CURRENT APPLICATION NUMBER: US/10/006,773
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: 60/250,089
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(410)
; OTHER INFORMATION: 3D8 Light chain V region, plus leader
US-10-006-773-10

Query Match      18.4%; Score 80; DB 13; Length 729;
Best Local Similarity 100.0%; Pred. No. 3.8e-34;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 347 ACACGTTTCGGAGGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGTAT 406
DB 373 ACACGTTTCGGAGGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGTAT 432
QY 407 CCATCTTCCACCATCCAGT 426
DB 433 CCATCTTCCACCATCCAGT 452

RESULT 12
US-08-779-784-16
; Sequence 16, Application US/08779784
; Publication No. US20020164325A1
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Moses
; APPLICANT: Miller, David J.
; APPLICANT: Asakura, Kunihiko
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,784
```

;; FILING DATE: 07-JAN-1997  
;; CLASSIFICATION: 424  
;; PRIORITY APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/692,084  
;; FILING DATE: 08-AUG-1996  
;; PRIORITY APPLICATION DATA: US 08/236,520  
;; APPLICATION NUMBER: US 08/236,520  
;; FILING DATE: 29-APR-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Jackson Esq., David A.  
;; REGISTRATION NUMBER: 26,742  
;; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201-487-5800  
;; TELEFAX: 201-343-1684  
;; INFORMATION FOR SEQ ID NO: 16:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 351 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; HYPOTHETICAL: NO  
US-08-779-784-16

Query Match 18.2%; Score 79; DB 8; Length 351;  
Best Local Similarity 100.0%; Pred. No. 1.3e-33;  
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 CCCAAGGTCAGTCGCGTAGTCTGGGTCAGATTATTTCTCACCATCAGCAGCCTT 294  
Db 241 CCCAAGGTCAGTCGCGTAGTCTGGGTCAGATTATTTCTCACCATCAGCAGCCTT 300

QY 295 GAGTCTGAAGATTTTGTAG 313  
Db 301 GAGTCTGAAGATTTTGTAG 319

RESULT 13  
US-09-840-459-95  
; Sequence 95, Application US/09840459  
; Patent No. US20020150576A1  
; GENERAL INFORMATION:  
; APPLICANT: Larosa, Gregory J.  
; APPLICANT: Horvath, Christopher  
; APPLICANT: Newman, Walter  
; APPLICANT: Jones, S. Tarran  
; APPLICANT: O'Brien, Slobhan H.  
; APPLICANT: O'Keefe, Theresa  
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
; FILE REFERENCE: 1855.1052-012  
; CURRENT APPLICATION NUMBER: US/09/840,459  
; CURRENT FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: PCT/US01/03537  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 09/497,625  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 09/359,193  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 09/121,781  
; PRIOR FILING DATE: 1998-07-23  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 95  
; LENGTH: 426  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-840-459-95

Query Match 18.2%; Score 79; DB 10; Length 426;  
Best Local Similarity 100.0%; Pred. No. 1.4e-33;  
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 CCCAAGGTCAGTCGCGTAGTCTGGGTCAGATTATTTCTCACCATCAGCAGCCTT 294  
Db 241 CCCAAGGTCAGTCGCGTAGTCTGGGTCAGATTATTTCTCACCATCAGCAGCCTT 300

QY 295 GAGTCTGAAGATTTTGTAG 313  
Db 301 GAGTCTGAAGATTTTGTAG 319

QY 342 TCCGTACAGTTCGGAGGGGGGACCAAGCTGGAAATAAAACGGGCTGATCTGCACCAAC 401  
Db 348 TCCGTACAGTTCGGAGGGGGGACCAAGCTGGAAATAAAACGGGCTGATCTGCACCAAC 407

QY 402 TGTATCCATCTTCCACCA 420  
Db 408 TGTATCCATCTTCCACCA 426

RESULT 14  
US-09-840-459-101/C  
; Sequence 101, Application US/09840459  
; Patent No. US20020150576A1  
; GENERAL INFORMATION:  
; APPLICANT: Larosa, Gregory J.  
; APPLICANT: Horvath, Christopher  
; APPLICANT: Newman, Walter  
; APPLICANT: Jones, S. Tarran  
; APPLICANT: O'Brien, Slobhan H.  
; APPLICANT: O'Keefe, Theresa  
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND  
; FILE REFERENCE: 1855.1052-012  
; CURRENT APPLICATION NUMBER: US/09/840,459  
; CURRENT FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: PCT/US01/03537  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 09/497,625  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 09/359,193  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 09/121,781  
; PRIOR FILING DATE: 1998-07-23  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 101  
; LENGTH: 426  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-840-459-101

Query Match 18.2%; Score 79; DB 10; Length 426;  
Best Local Similarity 100.0%; Pred. No. 1.4e-33;  
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 TCCGTACAGTTCGGAGGGGGGACCAAGCTGGAAATAAAACGGGCTGATCTGCACCAAC 401  
Db 79 TCCGTACAGTTCGGAGGGGGGACCAAGCTGGAAATAAAACGGGCTGATCTGCACCAAC 20

QY 402 TGTATCCATCTTCCACCA 420  
Db 19 TGTATCCATCTTCCACCA 1

RESULT 15  
US-10-006-773-14  
; Sequence 14, Application US/10006773  
; Publication No. US20020132983A1  
; GENERAL INFORMATION:  
; APPLICANT: Junghans, Richard P.  
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor A  
; FILE REFERENCE: 003  
; CURRENT APPLICATION NUMBER: US/10/006,773  
; CURRENT FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: 60/250,089  
; PRIOR FILING DATE: 2000-11-30  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 504  
; TYPE: DNA  
; ORGANISM: Mus sp.

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(402)
; OTHER INFORMATION: 4D4 Light chain V region, plus leader
US-10-006-773-14

Query Match      17.7%; Score 77; DB 13; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.8e-32;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 350 CGTTCGGAGGGGGACCAAGCTGGAATRAAARCGGGCTGATGCTGCACCAACTGTATCCA 409
Db 368 CGTTCGGAGGGGGACCAAGCTGGAATRAAARCGGGCTGATGCTGCACCAACTGTATCCA 427
QY 410 TCTTCCCACCATCCAGT 426
Db 428 TCTTCCCACCATCCAGT 444
```

Search completed: August 30, 2003, 22:02:42  
Job time : 160.756 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 18:52:40 ; Search time 1513.76 Seconds  
(without alignments)  
6984.224 Million cell updates/sec

Title: US-08-836-455-1

Perfect score: 435

Sequence: 1 ATGGGGGCCCTGCTCAGAT.....CACCATCCAGTAAGCTTGGG 435

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_estbma:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pin:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	20.0	594	10	BE309592 601094848
2	85	19.8	805	12	BI454240 603170666
3	85	19.5	585	12	BM194777 L0700H12-
4	85	19.5	725	12	BG963055 602828068

5	85	19.5	891	10	BF579422
6	85	19.5	906	10	BF785914
7	85	19.5	926	13	BQ959057
8	85	19.5	935	11	BC031349
9	84	19.3	606	13	BQ922747
10	84	19.3	854	13	BQ947692
11	84	19.3	871	13	BQ956722
12	84	19.3	886	13	BQ940987
13	84	19.3	908	13	BQ524124
14	84	19.3	941	13	BQ523453
15	84	19.3	975	12	BG961850
16	83	19.1	750	12	BG965050
17	81	18.6	374	13	BY083003
18	81	18.6	630	10	BF138788
19	81	18.6	707	12	BI250555
20	79	18.2	712	12	BI100311
21	79	18.2	778	12	BG964076
22	79	18.2	798	12	BG968518
23	78	17.9	959	13	BQ939046
24	75	17.2	944	10	BF687485
25	75	17.2	967	10	BF687410
26	73	16.8	793	12	BG965736
27	67	15.4	532	12	BI104783
28	67	15.4	624	10	BE306691
29	67	15.4	685	10	BE369854
30	67	15.4	695	10	BE284224
31	67	15.4	755	12	BI150509
32	67	15.4	762	10	BF144806
33	67	15.4	772	10	BE285427
34	67	15.4	827	12	BI152061
35	67	15.4	865	12	BI659552
36	67	15.4	874	10	BG518527
37	67	15.4	876	12	BI107286
38	67	15.4	882	10	BF159226
39	67	15.4	918	10	BF135931
40	67	15.4	936	10	BF164906
41	67	15.4	997	12	BI107100
42	66	15.2	748	12	BG963884
43	64	14.7	354	13	BY085718
44	57	13.1	698	12	BI109045
45	57	13.1	849	10	BF583521

ALIGNMENTS

RESULT 1  
BE309592  
LOCUS  
DEFINITION  
601094848F1 NCI\_CGAP\_Mam5 Mus musculus cdna clone IMAGE:3489635 5', linear EST 26-OCT-2000  
mRNA sequence.  
ACCESSION  
BE309592  
VERSION  
BE309592.1  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
REFERENCE  
1 (bases 1 to 594)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS  
National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE  
Unpublished  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: L1AM8530 row: o column: 12  
High quality sequence stop: 591.

```

FEATURES
source
Location/Qualifiers
1. .594
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3489635"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam5"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT      158 a      149 c      143 g      144 t
ORIGIN

Query Match      20.0%; Score 87; DB 10; Length 594;
Best Local Similarity 100.0%; Pred. No. 6.6e-35;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 340 TCTCCGTACACGTTCCGAGGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCA 399
      |||||||
Db 326 TCTCCGTACACGTTCCGAGGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCA 385
      |||||||
QY 400 ACTGTATCCATCTTCCCACCATCCAGT 426
      |||||||
Db 386 ACTGTATCCATCTTCCCACCATCCAGT 412
      |||||||

RESULT 2
BI454240
LOCUS      603170666F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5250017 5',
          mRNA sequence.
DEFINITION
ACCESSION      BI454240.1 GI:152444896
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 805)
Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Klotz,E.,
Kelsoe,G., Hodess,R. and Ko,M.S.H.
Systematic Analyses of NIA Mouse Germinal Center B Cell cDNA
Library
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11631 row: d column: 18
High quality sequence stop: 790.
Location/Qualifiers
1. .805
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5250017"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam5"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT      230 a      209 c      186 g      180 t
ORIGIN

Query Match      19.8%; Score 86; DB 12; Length 805;
Best Local Similarity 100.0%; Pred. No. 2.5e-34;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 CTCCTACACGTTCCGAGGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCA 400
      |||||||
Db 356 CTCCTACACGTTCCGAGGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCA 415
      |||||||
QY 401 CTGTATCCATCTTCCCACCATCCAGT 426
      |||||||
Db 416 CTGTATCCATCTTCCCACCATCCAGT 441
      |||||||

RESULT 3
BM194777/c
LOCUS      BM194777 585 bp mRNA linear EST 30-JAN-2002
DEFINITION      musculus cDNA clone L0700H12 3', mRNA sequence.
ACCESSION      BM194777
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 585)
Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Klotz,E.,
Kelsoe,G., Hodess,R. and Ko,M.S.H.
Systematic Analyses of NIA Mouse Germinal Center B Cell cDNA
Library
Unpublished
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: L0700 row: H column: 12
Seq primer: -21M13 Forward
High quality sequence stop: 585
POLYA=Yes.
Location/Qualifiers
1. .585
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="niaEST:L0700H12-3"
/clone="L0700H12"
/tissue_type="Germinal Center B Cell"
/lab_host="DH10B"
/clone_lib="NIA Mouse Germinal Center B Cell cDNA Library"
/notes="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA
).FACS-sorted Germinal Center B cells were provided by
Drs. Richard Hodess, Emily Klotz (National Institute on
Aging and National Cancer Institute, USA) and Garnett
Kelsoe (Duke University, USA). Double-stranded cDNAs were
synthesized from 0.46 ug of total RNA with an Oligo(dT)
primer [Invitrogen]:
5'-pGACTAGTCTAGATCGGAGCGGCCCTTTTCTTTT-3'],
treated with T4 DNA polymerase, and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker LL-Sal3 (Ref. Development 127: 1737-1749 (2000
) [PMID: 1072549]), purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, cDNAs
were amplified by long-range high fidelity PCR using Ex
Taq polymerase (Takara) and purified by phenol/chloroform,

```

followed by Centricon 100 purification. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is 1.2 kb. The library was constructed by Yulan Piao (NIA)."

```

BASE COUNT      137 a 121 c 167 g 160 t
ORIGIN
Query Match      19.5%; Score 85; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.6e-34;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 TCCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 401
      |||||||
Db 572 TCCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 513

QY 402 TGTATCCATCTTCCACCATCCAGT 426
      |||||||
Db 512 TGTATCCATCTTCCACCATCCAGT 488

RESULT 4
LOCUS      BG963055
DEFINITION 602828068F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4982825 5',
      mRNA sequence.
ACCESSION  BG963055
VERSION     BG963055.1 GI:14350692
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 725)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M10986 row: g column: 18
High quality sequence stop: 719.
Location/Qualifiers
1. 725
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/clone="IMAGE:4982825"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT      195 a 182 c 174 g 174 t
ORIGIN
Query Match      19.5%; Score 85; DB 12; Length 725;
Best Local Similarity 100.0%; Pred. No. 8.1e-34;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 TCCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 401
      |||||||
Db 372 TCCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 431

```

```

QY 402 TGTATCCATCTTCCACCATCCAGT 426
      |||||||
Db 432 TGTATCCATCTTCCACCATCCAGT 456

RESULT 5
LOCUS      BF579422
DEFINITION 602093833F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4208144 5',
      mRNA sequence.
ACCESSION  BF579422
VERSION     BF579422.1 GI:11653134
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 891)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M9772 row: a column: 09
High quality sequence stop: 711.
Location/Qualifiers
1. 891
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Co24"
/clone="IMAGE:4208144"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT      246 a 226 c 204 g 215 t
ORIGIN
Query Match      19.5%; Score 85; DB 10; Length 891;
Best Local Similarity 100.0%; Pred. No. 8.6e-34;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 TCCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 401
      |||||||
Db 353 TCCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 412

QY 402 TGTATCCATCTTCCACCATCCAGT 426
      |||||||
Db 413 TGTATCCATCTTCCACCATCCAGT 437

RESULT 6
LOCUS      BF785914
DEFINITION 602112548F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4240762
      5', mRNA sequence.
ACCESSION  BF785914
VERSION     BF785914.1 GI:12090950
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: crapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9855 row: p column: 11
High quality sequence stop: 718.
Location/Qualifiers
1..906
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4240762"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Kid4"
/notes="Organ: Kidney; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT 243 a 232 c 237 g 194 t
ORIGIN
Query Match 19.5%; Score 85; DB 10; Length 906;
Best Local Similarity 100.0%; Pred. No. 8.6e-34;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 TCCGTACACGTCGGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 401
Db 371 TCCGTACACGTCGGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 430

QY 402 TGTATCCATCTTCCACCATCCAGT 426
Db 431 TGTATCCATCTTCCACCATCCAGT 455

RESULT 7
BQ959057 926 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_10049748 NCI_CGAP_Co24 Mus musculus cDNA clone
DEFINITION IMAGE:6479377 5', mRNA sequence.
ACCESSION BQ959057
VERSION BQ959057.1 GI:22374535
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: crapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM14024 row: h column: 02
High quality sequence stop: 692.
Location/Qualifiers
1..926
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4240762"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Kid4"
/notes="Organ: Kidney; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT 243 a 232 c 237 g 194 t
ORIGIN
Query Match 19.5%; Score 85; DB 10; Length 906;
Best Local Similarity 100.0%; Pred. No. 8.6e-34;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 TCCGTACACGTCGGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 401
Db 371 TCCGTACACGTCGGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 430

QY 402 TGTATCCATCTTCCACCATCCAGT 426
Db 431 TGTATCCATCTTCCACCATCCAGT 455

RESULT 8
BQ959057 935 bp mRNA linear HTC 03-JUN-2002
LOCUS AGENCOURT_10049748 NCI_CGAP_Co24 Mus musculus cDNA clone
DEFINITION IMAGE:6479377 5', mRNA sequence.
ACCESSION BQ959057
VERSION BQ959057.1 GI:21315064
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: crapbs-r@mail.nih.gov
Tissue Procurement: Marcello Bento Soares, Ph.D.
cDNA Library Preparation: Soares Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pavan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 67 Row: j Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein
This clone has the following problem: no cloning site /
microdeletion.
Location/Qualifiers
1..935

```

/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1514385"  
 /tissue\_type="Mammary gland, lactating mouse"  
 /clone\_lib="Soares\_mammary\_gland\_NMLMG"  
 /lab\_host="DH10B"  
 /note="vector: pT7T3D-Pac"  
 BASE COUNT 251 a 247 c 208 g 229 t  
 ORIGIN

Query Match 19.3%; Score 85; DB 11; Length 935;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-34;  
 Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 342 TCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 401  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 344 TCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAAC 403  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 402 TGTATCCATCTCCACCAATCCAGT 426  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 404 TGTATCCATCTCCACCAATCCAGT 428

RESULT 9  
 BQ922747  
 LOCUS  
 DEFINITION BQ922747 606 bp mRNA linear EST 20-AUG-2002  
 IMAGE:6395978 5', mRNA sequence.  
 ACCESSION BQ922747  
 VERSION BQ922747.1 GI:223337778  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)

ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 606)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLMI3891 row: i column: 03  
 High quality sequence stop: 605.

FEATURES  
 source  
 Location/Qualifiers  
 1..606  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6395978"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Co24"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.6 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 181 a 169 c 121 g 135 t  
 ORIGIN

Query Match 19.3%; Score 84; DB 13; Length 606;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-33;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 343 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACT 402  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 15 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACT 74  
 QY 403 GTATCCATCTCCACCAATCCAGT 426  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 75 GTATCCATCTCCACCAATCCAGT 98

RESULT 10  
 BQ947692  
 LOCUS  
 DEFINITION BQ947692 854 bp mRNA linear EST 21-AUG-2002  
 IMAGE:6441024 5', mRNA sequence.  
 ACCESSION BQ947692  
 VERSION BQ947692.1 GI:223631170  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)

ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 854)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLMI3963 row: j column: 01  
 High quality sequence stop: 671.

FEATURES  
 source  
 Location/Qualifiers  
 1..854  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N-3"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6441024"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP Mam2"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"  
 BASE COUNT 234 a 237 c 190 g 193 t  
 ORIGIN

Query Match 19.3%; Score 84; DB 13; Length 854;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-33;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 343 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACT 402  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 361 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACT 420  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 403 GTATCCATCTCCACCAATCCAGT 426  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 421 GTATCCATCTCCACCAATCCAGT 444

RESULT 11  
 BQ956722  
 LOCUS  
 DEFINITION BQ956722 871 bp mRNA linear EST 21-AUG-2002  
 IMAGE:6396481 5', mRNA sequence.  
 ACCESSION BQ956722  
 VERSION BQ956722.1 GI:22372200  
 KEYWORDS EST.

SOURCE  
ORGANISM Mus musculus (house mouse)

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
JOURNAL 1 (bases 1 to 871)  
COMMENT NIH-MGC <http://mgc.nci.nih.gov/>.  
Unpublished National Institutes of Health, Mammalian Gene Collection (MGC)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: Agencourt Bioscience Corporation  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM13892 row: n column: 02  
High quality sequence stop: 630.  
Location/Qualifiers  
1. .871  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6396481"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Co24"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 218 a 268 c 182 g 193 t 10 others  
ORIGIN

Query Match 19.3%; Score 84; DB 13; Length 871;  
Best Local Similarity 100.0%; Pred. No. 2.9e-33;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 343 CCGTACACGTTCCGAGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACT 402  
|||||  
Db 190 CCGTACACGTTCCGAGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACT 249  
|||||

QY 403 GTATCCATCTCCACCATCCAGT 426  
|||||

Db 250 GTATCCATCTCCACCATCCAGT 273  
|||||

RESULT 12  
LOCUS BQ940987 886 bp mRNA linear EST 21-AUG-2002  
DEFINITION AGENCOURT\_8881891 NCI\_CGAP\_Co24 Mus musculus cDNA clone  
IMAGE:6476176 5', mRNA sequence.  
ACCESSION BQ940987  
VERSION BQ940987.1 GI:22356465  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 886)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
Unpublished National Institutes of Health, Mammalian Gene Collection (MGC)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: Agencourt Bioscience Corporation  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
JOURNAL 1 (bases 1 to 886)  
COMMENT NIH-MGC <http://mgc.nci.nih.gov/>.  
Unpublished National Institutes of Health, Mammalian Gene Collection (MGC)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: Agencourt Bioscience Corporation  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLAM14016 row: b column: 17  
High quality sequence stop: 626.  
Location/Qualifiers  
1. .886  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6476176"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Co24"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 242 a 243 c 183 g 216 t 2 others  
ORIGIN

Query Match 19.3%; Score 84; DB 13; Length 886;  
Best Local Similarity 100.0%; Pred. No. 2.9e-33;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 343 CCGTACACGTTCCGAGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACT 402  
|||||  
Db 316 CCGTACACGTTCCGAGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACT 375  
|||||

QY 403 GTATCCATCTCCACCATCCAGT 426  
|||||

Db 376 GTATCCATCTCCACCATCCAGT 399  
|||||

RESULT 13  
LOCUS BU524124 908 bp mRNA linear EST 13-SEP-2002  
DEFINITION AGENCOURT\_10126425 NCI\_CGAP\_Co24 Mus musculus cDNA clone  
IMAGE:6530882 5', mRNA sequence.  
ACCESSION BU524124  
VERSION BU524124.1 GI:22834563  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 908)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
Unpublished National Institutes of Health, Mammalian Gene Collection (MGC)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: Agencourt Bioscience Corporation  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM14133 row: f column: 02  
High quality sequence stop: 604.  
Location/Qualifiers  
1. .908  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6530882"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Co24"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 238 a 270 c 194 g 206 t  
ORIGIN

```

Query Match      19.3%; Score 84; DB 13; Length 908;
Best Local Similarity 100.0%; Pred. No. 2.9e-33;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 343 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACT 402
Db 173 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACT 232
QY 403 GTATCCATCTTCCACCATCCAGT 426
Db 233 GTATCCATCTTCCACCATCCAGT 256

RESULT 14
BU523453      941 bp      mRNA      linear      EST 13-SEP-2002
LOCUS      AGENCOURT_10169180 NCI_CGAP_Co24 Mus musculus cDNA clone
DEFINITION      IMAGE:6530124 5', mRNA sequence.
ACCESSION      BU523453
VERSION      BU523453.1 GI:22833892
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 941)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgabs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: Agencourt Bioscience Corporation
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM14131 row: f column: 12
            High quality sequence stop: 646.
            Location/Qualifiers
            1..941
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:6530124"
            /lab_host="NCI_CGAP_Co24"
            /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.6 kb. Constructed by Life
            Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT      249 a 261 c 214 g 217 t
ORIGIN

Query Match      19.3%; Score 84; DB 13; Length 941;
Best Local Similarity 100.0%; Pred. No. 3e-33;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 343 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACT 402
Db 380 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACT 439
QY 403 GTATCCATCTTCCACCATCCAGT 426
Db 440 GTATCCATCTTCCACCATCCAGT 463

RESULT 15
BG961850
LOCUS      BG961850      975 bp      mRNA      linear      EST 12-JUN-2001

```

```

DEFINITION      602826515F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4981443 5',
mRNA sequence.
ACCESSION      BC961850
VERSION      BC961850.1 GI:14349487
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 975)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgabs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: Incyte Genomics, Inc.
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM10982 row: n column: 04
            High quality sequence stop: 743.
            Location/Qualifiers
            1..975
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:4981443"
            /lab_host="DH10B (T1 phage-resistant)"
            /clone_lib="NCI_CGAP_Co24"
            /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.6 kb. Constructed by Life
            Technologies. Note: this is a NCI_CGAP Library."

```

```

FEATURES
source
BASE COUNT      288 a 264 c 210 g 213 t
ORIGIN

```

```

Query Match      19.3%; Score 84; DB 12; Length 975;
Best Local Similarity 100.0%; Pred. No. 3e-33;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 343 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACT 402
Db 363 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACT 422
QY 403 GTATCCATCTTCCACCATCCAGT 426
Db 423 GTATCCATCTTCCACCATCCAGT 446

Search completed: August 30, 2003, 21:57:07
Job time : 1518.76 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 13:43:39 ; Search time 1971.09 seconds  
(without alignments)  
9028.331 Million cell updates/sec

Title: US-08-836-455-1  
Perfect score: 435  
Sequence: 1 ATGGGGGCCCTCTCAGAT.....CACCATCCAGTAAGCTTGGG 435

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*\*

- 1: gb\_ba:\*\*
- 2: gb\_htg:\*\*
- 3: gb\_in:\*\*
- 4: gb\_cm:\*\*
- 5: gb\_ov:\*\*
- 6: gb\_pat:\*\*
- 7: gb\_ph:\*\*
- 8: gb\_pl:\*\*
- 9: gb\_pr:\*\*
- 10: gb\_ro:\*\*
- 11: gb\_sts:\*\*
- 12: gb\_sy:\*\*
- 13: gb\_un:\*\*
- 14: gb\_vl:\*\*
- 15: em\_ba:\*\*
- 16: em\_fun:\*\*
- 17: em\_hum:\*\*
- 18: em\_in:\*\*
- 19: em\_mu:\*\*
- 20: em\_or:\*\*
- 21: em\_ov:\*\*
- 22: em\_pat:\*\*
- 23: em\_ph:\*\*
- 24: em\_pi:\*\*
- 25: em\_ro:\*\*
- 26: em\_sts:\*\*
- 27: em\_un:\*\*
- 28: em\_vl:\*\*
- 29: em\_vl:\*\*
- 30: em\_htg\_hum:\*\*
- 31: em\_htg\_inv:\*\*
- 32: em\_htg\_other:\*\*
- 33: em\_htg\_mus:\*\*
- 34: em\_htg\_pln:\*\*
- 35: em\_htg\_rod:\*\*
- 36: em\_htg\_mam:\*\*
- 37: em\_htg\_vrt:\*\*
- 38: em\_sy:\*\*
- 39: em\_htgo\_hum:\*\*
- 40: em\_htgo\_mus:\*\*
- 41: em\_htgo\_other:\*\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	435	100.0	435	6	AR164505	AR164505 Sequence
2	435	100.0	435	6	BD085737	BD085737 Methods o
3	384.6	88.4	407	6	E54981	E54981 Peptide. 1/
4	379.4	87.2	381	10	AF124721	AF124721 Mus muscu
5	361.6	83.1	384	10	AB017434	AB017434 Mus muscu
6	361.6	83.1	390	10	MUSIKCC	L41880 Mus musculu
7	352.2	81.0	381	10	AF045508	AF045508 Mus muscu
8	349	80.2	381	10	AF045495	AF045495 Mus muscu
9	345.8	79.5	381	10	AF045510	AF045510 Mus muscu
10	332	76.3	380	10	MMIGGVJ1	X02177 M.musculus
11	327.2	75.2	381	10	MMIGGVJ2	X02178 M.musculus
12	325.4	74.8	383	10	MUSIGKCM	M12191 Mouse Ig ac
13	323.8	74.4	405	10	AB016620	AB016620 Mus muscu
14	323.4	74.3	387	6	AR169918	AR169918 Sequence
15	317.2	72.9	354	10	AB089681	AB089681 Mus muscu
16	316	72.6	684	10	MUSIGKAC1	J00565 Mus musculu
17	314.4	72.3	685	10	MMIGK7	V00808 Part of the
18	311.6	71.6	348	10	MMVJIG	X54755 Mouse rear
19	309.8	71.2	321	10	AF163749	AF163749 Mus muscu
20	309.4	71.1	739	6	AR007981	AR007981 Sequence
21	309.4	71.1	739	6	AR058996	AR058996 Sequence
22	309.4	71.1	739	6	I23446	I23446 Sequence 5
23	308.6	70.9	324	6	AX722008	AX722008 Sequence
24	308.4	70.9	729	6	E54976	E54976 Peptide. 1/
25	307.6	70.7	324	10	MUSX	L48667 Mus musculu
26	307.4	70.7	711	6	E54975	E54975 Peptide. 1/
27	306.4	70.4	324	6	E54967	E54967 Peptide. 1/
28	305.4	70.2	328	10	MMU55591	U55591 Mus musculu
29	301.4	69.3	756	6	AX256284	AX256284 Sequence
30	301.4	69.3	771	6	AX256296	AX256296 Sequence
31	301.4	69.3	1497	6	AX256302	AX256302 Sequence
32	300.6	69.1	324	6	I03643	I03643 Sequence 4
33	300.6	69.1	324	6	I07835	I07835 Sequence 4
34	299	68.7	323	10	AY229938	AY229938 Mus muscu
35	298.6	68.6	321	10	MUSIGKAA3	M59920 Mouse IG ge
36	298.4	68.6	348	10	AY245603	AY245603 Mus muscu
37	297.2	68.3	1019	10	BC027418	BC027418 Mus muscu
38	296.2	68.1	413	10	MUSIGKCLN	M20832 Mouse Igmk
39	295.6	68.0	972	10	AF466770	AF466770 Mus muscu
40	292.4	67.2	959	10	BC015292	BC015292 Mus muscu
41	292.2	67.2	321	6	AR081902	AR081902 Sequence
42	292.2	67.2	321	6	AR081905	AR081905 Sequence
43	291.6	67.0	326	10	MMVJIG3	X54756 Mouse rear
44	291.2	66.9	321	6	AR081908	AR081908 Sequence
45	291.2	66.9	323	12	AY008757	AY008757 Synthetic

ALIGNMENTS

RESULT 1	AR164505	AR164505	435 bp	DNA	linear	PAT 17-OCT-2001
LOCUS	AR164505	Sequence. 1 from patent US 6274143.				
DEFINITION	AR164505					
ACCESSION	AR164505					
VERSION	AR164505.1	GI:16237555				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1. (bases 1 to 435)					
AUTHORS	Chatterjee, M. and Foon, K.A.					
TITLE	Methods of delaying development of HMFG-associated tumors using anti-idiotype antibody 11D10					
JOURNAL	Patent: US 6274143-A 1 14-AUG-2001;					

```
FEATURES          Location/Qualifiers
source            1..435
BASE COUNT       100 a 111 c 102 g 122 t
ORIGIN
Query Match      100.0%; Score 435; DB 6; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.7e-120;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGGCCCCCTGCTCAGATCTTGGGTTCTTGTGCTTCTTGTGCTTCCAGGTACCAGATGT 60
Db 1 ATGGGGCCCCCTGCTCAGATCTTGGGTTCTTGTGCTTCTTGTGCTTCCAGGTACCAGATGT 60
QY 61 GACATCCAGATGACCCAGTCTCCATCCCTTATCTGCTCTCTGCGACAAAGAGTCACT 120
Db 61 GACATCCAGATGACCCAGTCTCCATCCCTTATCTGCTCTCTGCGACAAAGAGTCACT 120
QY 121 CTCACCTTGTGCGGCAAGTCAGGACATGGTATTAACTTACATTCAGTTCAGCAGGAACCA 180
Db 121 CTCACCTTGTGCGGCAAGTCAGGACATGGTATTAACTTACATTCAGTTCAGCAGGAACCA 180
QY 181 GATGGAACCTATTAAACGCTGATCTACGCCATCCAGTTTAGCTTCTGCTGCCCAAA 240
Db 181 GATGGAACCTATTAAACGCTGATCTACGCCATCCAGTTTAGCTTCTGCTGCCCAAA 240
QY 241 AGGTTTCAGTGGCAGTGTGGTCTGAGATTTATCTTCACCATCAGCAGCCTTGAAGTCT 300
Db 241 AGGTTTCAGTGGCAGTGTGGTCTGAGATTTATCTTCACCATCAGCAGCCTTGAAGTCT 300
QY 301 GAAGATTTGTAGGCTTACTGCTACAAATATGCTAGTTCCTCGTACACGTTTCGGAGGG 360
Db 301 GAAGATTTGTAGGCTTACTGCTACAAATATGCTAGTTCCTCGTACACGTTTCGGAGGG 360
QY 361 GGGACCAAGCTGGAATAAAGCGGCTGATCTGCAACATGCTATCCATCTTCCCAACCA 420
Db 361 GGGACCAAGCTGGAATAAAGCGGCTGATCTGCAACATGCTATCCATCTTCCCAACCA 420
QY 421 TCCAGTAAGCTTGGG 435
Db 421 TCCAGTAAGCTTGGG 435

RESULT 2
BD085737
LOCUS
DEFINITION
  Methods of delaying development of HMFG-associated tumors using
  anti-idiotypic antibody 11D10.
ACCESSION
  BD085737
VERSION
  JP 2001523269-A/1
KEYWORDS
  unidentified
SOURCE
  unclassified.
ORGANISM
  1 (bases 1 to 435)
REFERENCE
  Chatterjee,M. and Poon,K.A.
AUTHORS
  Methods of delaying development of HMFG-associated tumors using
  anti-idiotypic antibody 11D10
TITLE
  Patent: JP 2001523269-A 1 20-NOV-2001;
JOURNAL
  THE UNIVERSITY OF KENTUCKY RESEARCH FOUNDATION
COMMENT
  OS Unidentified
  PN JP 2001523269-A/1
  PD 20-NOV-2001
  PR 12-JUN-1998 JP 1999503252
  PF 13-JUN-1997 US 60/049540,11-JUN-1998 US 09/096244 PI
  MALAYA CHATTERJEE,KENNETH A POON
  PC A61K39/395,A61K39/39//C07K16/42
  CC Strandedness: Single;
  CC Topology: Linear;
  CC Methods of delaying development of HMFG-associated tumors CC
  using
  CC anti-idiotypic antibody 11D10
  FH Key Location/Qualifiers

FEATURES          Location/Qualifiers
source            1..435
BASE COUNT       100 a 111 c 102 g 122 t
ORIGIN
Query Match      100.0%; Score 435; DB 6; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.7e-120;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGGCCCCCTGCTCAGATCTTGGGTTCTTGTGCTTCTTGTGCTTCCAGGTACCAGATGT 60
Db 1 ATGGGGCCCCCTGCTCAGATCTTGGGTTCTTGTGCTTCTTGTGCTTCCAGGTACCAGATGT 60
QY 61 GACATCCAGATGACCCAGTCTCCATCCCTTATCTGCTCTCTGCGACAAAGAGTCACT 120
Db 61 GACATCCAGATGACCCAGTCTCCATCCCTTATCTGCTCTCTGCGACAAAGAGTCACT 120
QY 121 CTCACCTTGTGCGGCAAGTCAGGACATGGTATTAACTTACATTCAGTTCAGCAGGAACCA 180
Db 121 CTCACCTTGTGCGGCAAGTCAGGACATGGTATTAACTTACATTCAGTTCAGCAGGAACCA 180
QY 181 GATGGAACCTATTAAACGCTGATCTACGCCATCCAGTTTAGCTTCTGCTGCCCAAA 240
Db 181 GATGGAACCTATTAAACGCTGATCTACGCCATCCAGTTTAGCTTCTGCTGCCCAAA 240
QY 241 AGGTTTCAGTGGCAGTGTGGTCTGAGATTTATCTTCACCATCAGCAGCCTTGAAGTCT 300
Db 241 AGGTTTCAGTGGCAGTGTGGTCTGAGATTTATCTTCACCATCAGCAGCCTTGAAGTCT 300
QY 301 GAAGATTTGTAGGCTTACTGCTACAAATATGCTAGTTCCTCGTACACGTTTCGGAGGG 360
Db 301 GAAGATTTGTAGGCTTACTGCTACAAATATGCTAGTTCCTCGTACACGTTTCGGAGGG 360
QY 361 GGGACCAAGCTGGAATAAAGCGGCTGATCTGCAACATGCTATCCATCTTCCCAACCA 420
Db 361 GGGACCAAGCTGGAATAAAGCGGCTGATCTGCAACATGCTATCCATCTTCCCAACCA 420
QY 421 TCCAGTAAGCTTGGG 435
Db 421 TCCAGTAAGCTTGGG 435

RESULT 3
E54981
LOCUS
DEFINITION
  Peptide.
ACCESSION
  E54981
VERSION
  E54981.1 GI:18629719
KEYWORDS
  JP 2000236884-A/15.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
REFERENCE
  1 (bases 1 to 407)
AUTHORS
  Nishida,T., Okura,T., Tanimoto,T. and Kurimoto,M.
TITLE
  Peptide
JOURNAL
  Patent: JP 2000236884-A 15 05-SEP-2000;
COMMENT
  OS Mus musculus (mouse)
  PN JP 2000236884-A/15
  PD 05-SEP-2000
  PF 24-JUN-1999 JP 1999177846
  PR
  PI TAKEHIRO NISHIDA,TAKANORI OKURA,TADAO TANIMOTO,PI MASASHI
  KURIMOTO
  PC C12N15/09,A61K31/00,A61K39/395,A61K48/00,C07K16/24,C12P21/08,
  CC C12N15/00
```





```

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 381)
O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.
Anti-DNA antibodies of normal mice immunized with poly(dC) are
structurally similar to natural autoantibodies
Unpublished
2 (bases 1 to 381)
O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.
Submitted (02-FEB-1998) Biochemistry, Tufts University School of
Medicine, 136 Harrison Avenue, Boston, MA 02111, USA
FEATURES
    source
        location/Qualifiers
            1..381
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6"
                /db_xref="taxon:10090"
                /chromosome="6"
                /clone="dc10"
                /cell_line="spleen cell hybridoma"
                /gene="Igk"
                /gene="Igk"
                /codon_start=1
                /product="anti-poly(dC) monoclonal antibody kappa light
                chain"
                /protein_id="AAC04536.1"
                /db_xref="GI:2906100"
                /translation="MRAPQIFGFLLLFPGTRCDIOMTQSPSLSASLGRVSLTCR
                ASDIGSLNLWQEPDGTIKRLIYATSSLDGVPKRFSGRSGDYSLTISSESD
                FVDYICLYASFPWTFGGTKLEIK"
sig_peptide
    1..60
V_region
    61..345
J_segment
    346..>381
BASE COUNT      91 a    90 g    109 t
ORIGIN
Query Match      81.08; Score 352.2; DB 10; Length 381;
Best Local Similarity 93.3; Pred. No. 2e-95;
Matches 363; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
Qy  1 ATGGGGCCCTGCTCAGATTCTTGGGTTCTTGTGCTCTGTGCTTCTCCAGGTACCAAGTGT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1 ATGAGGGCTCTCGACAGATTTTGGCTTCTTGTGCTCTGTGCTTCTCCAGGTACCAAGTGT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  61 GACATCCAGATGACCCAGTCTCCATCCTCTTATCTGCTCTCTGCGTCTCTGGGACAAAGAGTCA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  61 GACATCCAGATGACCCAGTCTCCATCCTCTTATCTGCTCTCTGCGTCTCTGGGACAAAGAGTCA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTACCTTCCAGTACCAAGACCA 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTACCTTCCAGTACCAAGACCA 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  181 GATGGAACCTATTAACGCCCTGATCTACGCCACATCCAGTTAGTCTTCTGGTGTCGCCAAA 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  181 GATGGAACCTATTAACGCCCTGATCTACGCCACATCCAGTTAGTCTTCTGGTGTCGCCAAA 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  241 AGGTTTCAGTGGCAGTAGTCTGGGTTCAGATTATTTCTCTCACCATCAGCAGCCTTGGTCT 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  241 AGGTTTCAGTGGCAGTAGTCTGGGTTCAGATTATTTCTCTCACCATCAGCAGCCTTGGTCT 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  301 GAAGATTTTGTAGCCTATTACTGTCTACAATATGCTAGTCTCCGTACACGTTCCGAGGG 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  301 GAAGATTTTGTAGCCTATTACTGTCTACAATATGCTAGTCTCCGTACACGTTCCGAGGG 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  361 GGGACCAAGCTGGGAATAAA 381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  361 GGCACCAAGCTGGGAATAAA 381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 8
AF045495
LOCUS      381 bp      mRNA      linear      ROD 28-FEB-1998
DEFINITION Mus musculus dC4 anti-poly(dC) monoclonal antibody kappa light
chain variable region, (Igk) mRNA, partial cds.
ACCESSION AF045495
VERSION    AF045495.1 GI:2906073
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 381)
AUTHORS   O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.
TITLE     Anti-DNA antibodies of normal mice immunized with poly(dC) are
structurally similar to natural autoantibodies
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 381)
AUTHORS   O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.
TITLE     Direct Submission
JOURNAL   Submitted (02-FEB-1998) Biochemistry, Tufts University School of
Medicine, 136 Harrison Avenue, Boston, MA 02111, USA
FEATURES   Location/Qualifiers
            1..381
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6"
                /db_xref="taxon:10090"
                /chromosome="6"
                /clone="dc4"
                /cell_line="spleen cell hybridoma"
                /gene="Igk"
                /gene="Igk"
                /codon_start=1
                /product="anti-poly(dC) monoclonal antibody kappa light
                chain"
                /protein_id="AAC04523.1"
                /db_xref="GI:2906074"
                /translation="MRAPQIFGFLLLFPGTRCDIOMTQSPSLSASLGRVSLTCR
                ASDIGSLNLWQEPDGTIKRLIYATSSLDGVPKRFSGRSGDYSLTISSESD
                FVDYICLYASFPWTFGGTKLEIK"
sig_peptide
    1..60
V_region
    61..345
J_segment
    346..>381
BASE COUNT      93 a    89 g    109 t
ORIGIN
Query Match      80.28; Score 349; DB 10; Length 381;
Best Local Similarity 94.8; Pred. No. 1.9e-94;
Matches 361; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy  1 ATGGGGCCCTGCTCAGATTCTTGGGTTCTTGTGCTCTTGTGCTTCTCCAGGTACCAAGTGT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1 ATGAGGGCTCTCGACAGATTTTGGGTTCTTGTGCTCTTGTGCTTCTCCAGGTACCAAGTGT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  61 GACATCCAGATGACCCAGTCTCCATCCTCTTATCTGCTCTCTGGGACAAAGAGTCA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  61 GACATCCAGATGACCCAGTCTCCATCCTCTTATCTGCTCTCTGGGACAAAGAGTCA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTACCTTCCAGTACCAAGACCA 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTACCTTCCAGTACCAAGACCA 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  181 GATGGAACCTATTAACGCCCTGATCTACGCCACATCCAGTTAGTCTTCTGGTGTCGCCAAA 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  181 GATGGAACCTATTAACGCCCTGATCTACGCCACATCCAGTTAGTCTTCTGGTGTCGCCAAA 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

241	AGTTTCAGTGGCAGTAGTCTGGGT	CAGATTATTCTCTCACCATCAGCAGCCTTGAGTCT	300
241	AGTTTCAGTGGCAGTAGTCTGGGT	CAGATTATTCTCTCACCATCAGCAGCCTTGAGTCT	300
301	GAAGATTTTGTAGCCTATTACTGCT	TACAATATGCTAGTTCTCGGTACACGTTCCGAGGG	360
301	GAAGATTTTGTAGCCTATTACTGCT	TACAATATGCAAGTTTTCGGTGGACGTTCCGTGGGA	360
361	GGGACCAAGCTGGAAATAAAA	381	
361	GGCACAAGCTGGAAATCAAA	381	

RESULT 9  
 AF045510 LOCUS  
 DEFINITION  
 AF045510 381 bp mRNA linear ROD 28-FEB-1998  
 Mus musculus 6C9 monoclonal antibody kappa light chain variable  
 region, (Igh) mRNA, partial cds.  
 AF045510  
 AF045510.1 GI:2906103  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;  
 1 (bases 1 to 381)  
 O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.  
 Anti-DNA antibodies of normal mice immunized with poly(dC) are  
 structurally similar to natural autoantibodies  
 unpublished  
 2 (bases 1 to 381)  
 O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.  
 Direct Submission  
 Submitted (02-FEB-1998) Biochemistry, Tufts University School of  
 Medicine, 136 Harrison Avenue, Boston, MA 02111, USA  
 Location/Qualifiers  
 FEATURES

61	Db	 GACATCCAGATGACCCAGAGTCTCCATCCTCTTATCTGCTCTCTGGGAGAAAAGAGTCAGT	120
121	Qy	 CTCAGCTGTGCGGCAAGTCAGGACATTGGTATTAACTTTACATTTGGCTTCAGCAGGAACCA	180
121	Db	 CTCAGCTGTGCGGCAAGTCAGGACATTGGTAGAGCTTAACTGGCTTCAGCAGGAACCA	180
181	Qy	 GATGGAACATTATAACGCCCTGATCTACGCCACATCCAGTTTAGTTCTGGTGTCCCCAAA	240
181	Db	 GATGGAACATTATAACACCTGATCTACGCCACATCCAGTTTAGATTCTGGTGTCCCCAAA	240
241	Qy	 AGGTTTCAGTGGCAGTAGGTCTGGGTCCAGATTATTCTCTCACCATCAGCAGCCTTGAGTCT	300
241	Db	 AGGTTTCAGTGGCAGTAGGTCTGGGTCCAGATTATTCTCTCACCATCAGCAGCCTTGAGTCT	300
301	Qy	 GAAGATTTTGTAGCCTATTACTGCTACAAATATGCTAGTTCTCCGTACACGTTTCGGGAGGG	360
301	Db	 GAAGATTTTGTAGACTATTACTGCTACAAATATGCTAGTTTCCATTACAGTTTCGGCTCG	360
361	Qy	 GGGACCAAGCTGGAAATAAAA	381
361	Db	 GGGACAAAGTTGGAAATAAAA	381

RESULT 10	
MMIGGVJ1	
LOCUS	380 bp mRNA linear ROD 03-FEB-1995
DEFINITION	M.musculus mRNA for IgG kappa light chain (partial) Gloop 1.
ACCESSION	X02177
VERSION	X02177.1 GI:51894
KEYWORDS	gamma-immunoglobulin; Ig light chain; immunoglobulin; joining region; variable region.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE	
AUTHORS	Darley M.J. and Rees, A.R.
TITLE	Nucleotide sequences of five anti-lysozyme monoclonal antibodies
JOURNAL	EMBO J. 4 (2), 393-398 (1985)
MEDLINE	85257466
PUBMED	2410256









**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 16:05:49 ; Search time 190.798 Seconds  
(without alignments)  
6154.444 Million cell updates/sec

Title: US-08-836-455-1  
Perfect score: 435  
Sequence: 1 ATGGGGGCCCTGCTCAGAT.....CACCATCCAGTAAGCTTGGG 435

Scoring table: IDENTITY\_NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*

2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*

3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*

4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*

5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*

6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*

7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*

8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*

9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*

10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*

11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*

12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*

13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*

14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*

15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*

16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*

17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*

18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*

19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*

20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*

21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*

22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*

23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*

24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	435	100.0	435	18	Murine monoclonal
2	435	100.0	435	20	Antibody 11D10 lig
3	435	100.0	435	25	Mouse 11D10 antibo
4	394.8	90.8	450	4	Sequence encoding
5	384.6	88.4	407	21	Mouse light chain
6	323.4	74.3	387	18	Murine anti-human
7	323.4	74.3	387	24	Murine 44H104 mab
8	309.4	71.1	739	14	Sequence encoding

9	309.4	71.1	739	17	AAT36880	520C9 anti-c-erbB-
10	309.4	71.1	739	19	AAV21798	520C9 anti-c-erbB-
11	309.4	71.1	739	20	AAV63399	520C9 sfv DNA sequ
12	308.4	70.9	729	21	AA249543	PEscFv125-2H.Hf r
13	307.4	70.7	711	21	AA249542	PEscFv125-2H.Hf r
14	306.4	70.4	324	21	AA249534	Mouse anti-IL-18 a
15	302	69.4	1605	14	AAQ46086	Sequence encoding
16	301.4	69.3	756	24	AAQ97136	p4-3 single chain
17	301.4	69.3	771	24	AAQ97142	p5-10 single chain
18	301.4	69.3	1497	24	AAQ97145	3B10xP4-3 bispecif
19	300.2	69.0	321	21	AAA38908	520C9 hybridoma VL
20	292.2	67.2	321	16	AAQ97504	Light chain variab
21	292.2	67.2	321	16	AAQ97507	Light chain variab
22	291.2	66.9	321	16	AAQ97510	Light chain variab
23	284.4	65.4	432	12	AAQ15114	IL-2 chimeric anti
24	284.4	65.4	456	22	AAQ66996	Filamentous phage
25	282.8	65.0	432	12	AAQ15113	IL-2 chimeric anti
26	282.4	64.9	867	22	AAQ67002	Filamentous phage
27	268.4	61.7	438	22	AAH41157	Human coding seque
28	267.6	61.5	447	17	AAT34541	Monoclonal anti-id
29	267.6	61.5	447	17	AAT31540	3H1 light chain va
30	267.6	61.5	447	18	AAT99434	Anti-idiotypic anti
31	267.6	61.5	447	21	AA235842	Murine MAb against
32	267.6	61.5	447	22	AAQ09316	Mouse anti-idiotyp
33	267.6	61.5	447	24	ABL60799	Antibody 3H1 light
34	267.6	61.5	447	24	AAK98279	Mouse 3H1 antibody
35	266.8	61.3	1106	24	ABO54241	Human ovarian anti
36	266.6	61.3	321	21	AAA38909	650E2 hybridoma VL
37	266	61.1	447	25	AA51275	Mouse 3H1 antibody
38	265.2	61.0	449	15	AAQ78732	Murine anti-human
39	263.6	60.6	439	20	AA221995	Partial nucleotide
40	263.6	60.6	714	21	AAA46899	DNA encoding the k
41	263.2	60.5	276	20	AAQ00879	Mouse derived RT3
42	262	60.2	439	18	AAQ73443	Human immunoglobul
43	262	60.2	817	21	AAA27389	Human IGFAM-9 immu
44	262	60.2	979	24	AB552791	Murine m166 antibo
45	261.6	60.1	276	20	ABX00875	Mouse derived RT3

ALIGNMENTS

RESULT 1	AAAT85149	standard; cDNA; 435 BP.
ID	AAAT85149	
XX	AC	AAAT85149;
XX	AC	AAAT85149;
DT	25-MAR-2003	(updated)
DT	04-JAN-1998	(first entry)
XX	XX	Murine monoclonal anti-idiotypic antibody 11D10 VL cDNA.
XX	XX	Monoclonal antibody 11D10; anti-idiotypic antibody; mucin;
KW	human milk fat globule; HMF; tumour; breast cancer; vaccine; ss.	
XX	XX	Mus musculus.
OS	XX	
XX	XX	
PH	key	Location/Qualifiers
FT	sig_peptide	1..60
FT	mat_peptide	/*tag= a
FT	FT	61..435
FT	FT	/*tag= b
XX	XX	
PN	WO9722699-A2.	
XX	XX	
PD	26-JUN-1997.	
XX	XX	
PF	19-DEC-1996;	96WO-US20757.
XX	XX	
PR	20-DEC-1995;	95US-0575762.
PR	26-JAN-1996;	96US-0591965.
PR	13-DEC-1996;	96US-0766350.

XX (KENT ) UNIV KENTUCKY.  
XX Chatterjee M, Chatterjee SK, Foon KA;  
XX WPI; 1997-341690/31.  
XX P-PSDB; AAW87119.  
XX Monoclonal anti-idiotype antibody 11D10 - elicits immune response  
XX against human milk fat globule disease associated tumours,  
XX especially breast cancer  
XX Claim 11; Page 94; 130pp; English.  
XX This CDNA sequence encodes the light chain variable region VL  
XX (AAW85149) of monoclonal anti-idiotype antibody 11D10 produced by  
XX hybridoma cell line ATCC 12020. 11D10 was obtained by immunising  
XX naive mice with MC-10 anti-HMFG antibody to obtain an anti-idiotype  
XX response. It elicits an immune response against a specific epitope  
XX of a high mol.wt. mucin of human milk fat globule (HMFG). It  
XX induces an immunological response to HMFG in mice, rabbits, monkeys  
XX and patients with advanced HMFG-associated tumours. Pharmaceutical  
XX compositions and vaccines comprising 11D10, 11D10 polypeptides  
XX and/or 11D10 polynucleotides are claimed. Also claimed are  
XX diagnostic kits and methods of using 11D10, 11D10 polypeptides  
XX and/or 11D10 polynucleotides, including methods of treating HMFG-  
XX associated tumours.  
XX (Updated on 25-MAR-2003 to correct PR field.)  
XX Sequence 435 BP; 100 A; 111 C; 102 G; 122 T; 0 other;

Query Match 100.0%; Score 435; DB 18; Length 435;  
Best Local Similarity 100.0%; Pred. No. 6.8e-128;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGCCCTGCTCAGATCTTGGGTTCTTGTCTCTTGTCTTCCAGGTACCAGATGT 60  
DB 1 ATGGGGCCCTGCTCAGATCTTGGGTTCTTGTCTCTTGTCTTCCAGGTACCAGATGT 60  
QY 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTCTGGGACAAAGAGTCAGT 120  
DB 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTCTGGGACAAAGAGTCAGT 120  
QY 121 CTCACCTGTGCGGCAAGTCAGGACATGGTATTAACCTTACATTCGCTTCAGCAGAACCA 180  
DB 121 CTCACCTGTGCGGCAAGTCAGGACATGGTATTAACCTTACATTCGCTTCAGCAGAACCA 180  
QY 181 GATGGAACCTATTAACGCGCTGATCTAGCCACATCCAGTTTGTAGTCTCTGGTGCCCAAA 240  
DB 181 GATGGAACCTATTAACGCGCTGATCTAGCCACATCCAGTTTGTAGTCTCTGGTGCCCAAA 240  
QY 241 AGTTTCAGTGGCAGTAGTCTGGGTCTGATTAATCTCTCACCATCAGCAGCCTTGTAGTCT 300  
DB 241 AGTTTCAGTGGCAGTAGTCTGGGTCTGATTAATCTCTCACCATCAGCAGCCTTGTAGTCT 300  
QY 301 GAACATTTGTAGCCTATTACTGCTACATATGCTAGTCTCCGTACAGTTCGGAGGG 360  
DB 301 GAAGATTTGTAGCCTATTACTGCTACATATGCTAGTCTCCGTACAGTTCGGAGGG 360  
QY 361 GGGACCAAGCTGGAATAAAGCGGCTGATCTGCACCAACTGATCCATCTCCCAACCA 420  
DB 361 GGGACCAAGCTGGAATAAAGCGGCTGATCTGCACCAACTGATCCATCTCCCAACCA 420  
QY 421 TCCAGTAAGCTTGGG 435  
DB 421 TCCAGTAAGCTTGGG 435

RESULT 2  
AAV83772  
ID AAV83772 standard; cDNA; 435 BP.  
XX  
AC AAV83772;

XX 16-MAR-1999 (first entry)  
XX Anticbody 11D10 light chain variable region coding sequence.  
XX Murine; mouse; antibody; light chain; variable region; anti-idiotype; ss;  
XX human milk fat globule; tumour; ovary; lung; pancreas; carcinoma; breast.  
XX Mus sp.  
XX Key Location/Qualifiers  
XX CDS 1..435  
XX /tag= a  
XX /product= "antibody 11D10 light chain variable region"  
XX /transl\_except= (pos:163..165, aa:Thr)  
XX /note= "no stop codon is given at the 3' end of the  
XX sequence"  
XX sig\_peptide 1..60  
XX mat\_peptide 61..435  
XX /tag= c  
XX WO9856419-A1.  
XX 17-DEC-1998.  
XX 12-JUN-1998; 98WO-US12250.  
XX 11-JUN-1998; 98US-0096244.  
XX 13-JUN-1997; 97US-0049540.  
XX (KENT ) UNIV KENTUCKY RES FOUND.  
XX Chatterjee M, Foon KA;  
XX WPI; 1999-060029/05.  
XX P-PSDB; AAW87593.  
XX Delaying development of, or treating, HMFG-associated tumours -  
XX using anti-idiotype antibody 11D10 raised against antibodies to  
XX human milk fat globule protein  
XX Disclosure; Fig 1; 54pp; English.  
XX This sequence represents the coding sequence for the murine antibody  
XX 11D10 light chain variable region. This anti-idiotype antibody is used  
XX to delay the development of, or treat, a human milk fat globule (HMFG)  
XX associated tumour in an individual having low tumour burden.  
XX The antibody 11D10 is used to prevent the recurrence of HMFG-associated  
XX tumours e.g. ovarian, non-small cell lung and pancreatic carcinoma,  
XX especially for treating breast tumours.  
XX Sequence 435 BP; 100 A; 111 C; 102 G; 122 T; 0 other;

Query Match 100.0%; Score 435; DB 20; Length 435;  
Best Local Similarity 100.0%; Pred. No. 6.8e-128;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGCCCTGCTCAGATCTTGGGTTCTTGTCTCTTGTCTTCCAGGTACCAGATGT 60  
DB 1 ATGGGGCCCTGCTCAGATCTTGGGTTCTTGTCTCTTGTCTTCCAGGTACCAGATGT 60  
QY 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTCTGGGACAAAGAGTCAGT 120  
DB 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTCTGGGACAAAGAGTCAGT 120  
QY 121 CTCACCTGTGCGGCAAGTCAGGACATGGTATTAACCTTACATTCGCTTCAGCAGAACCA 180  
DB 121 CTCACCTGTGCGGCAAGTCAGGACATGGTATTAACCTTACATTCGCTTCAGCAGAACCA 180  
QY 181 GATGGAACCTATTAACGCGCTGATCTAGCCACATCCAGTTTGTAGTCTCTGGTGCCCAAA 240  
DB 181 GATGGAACCTATTAACGCGCTGATCTAGCCACATCCAGTTTGTAGTCTCTGGTGCCCAAA 240

QY 241 AGGTTGAGTGGCAGTAGGTTGGTCTAGATTTATCTCTCACCATCAGCAGCCTTGAGTCT 300  
 |||||  
 DB 241 AGGTTGAGTGGCAGTAGGTTGGTCTAGATTTATCTCTCACCATCAGCAGCCTTGAGTCT 300  
 |||||  
 QY 301 GAAGATTTGTAGCCTTATCTCTCTACAAATATGCTAGTCTCCGTACAGGTTCCGAGGG 360  
 |||||  
 DB 301 GAAGATTTGTAGCCTTATCTCTCTACAAATATGCTAGTCTCCGTACAGGTTCCGAGGG 360  
 |||||  
 QY 361 GGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
 |||||  
 DB 361 GGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
 |||||  
 QY 421 TCCAGTAAGCTTGGG 435  
 |||||  
 DB 421 TCCAGTAAGCTTGGG 435  
 |||||

RESULT 3  
 AAL51273  
 ID AAL51273 standard; cDNA; 435 BP.  
 XX  
 AC AAL51273;  
 AC  
 DT 20-MAR-2003 (first entry)  
 XX  
 DE Mouse 11D10 antibody light chain variable region coding sequence.  
 XX  
 KW Mouse; murine; gene; ss; vaccine; tumour; human milk fat globules; HMFG;  
 KW carcinoembryonic antigen; CEA; 11D10; 3H1; HMFG-associated tumour;  
 KW CEA-associated tumour; anti-idiotypic antibody.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..435  
 FT /\*tag= a  
 FT /partial  
 FT /product= "Mouse 11D10 anti-idiotypic antibody light chain  
 FT variable region"  
 FT /note= "No stop codon is given"  
 FT sig\_peptide 1..60  
 FT /\*tag= b  
 FT mat\_peptide 61..435  
 FT /\*tag= c  
 FT  
 XX WO200292012-A2.  
 PN  
 XX  
 PD 21-NOV-2002.  
 XX  
 PF 17-MAY-2002; 2002WO-US15840.  
 XX  
 PR 17-MAY-2001; 2001US-0861294.  
 XX  
 PA (KENT ) UNIV KENTUCKY RES FOUND.  
 XX  
 PI Chatterjee M, Foon KA;  
 XX  
 DR WPI; 2003-129216/12.  
 DR P-PSDB; AAO16292.  
 XX  
 XX Use of anti-idiotypic antibodies for human milk fat globules (HMFG) - or  
 FT carcinoembryonic antigen (CEA)-associated tumor for delaying the  
 PT development of, or treating a HMFG- or CEA-associated tumor (e.g.  
 PT breast tumor) in humans  
 XX  
 PS Disclosure; Fig 1; 98pp; English.  
 XX  
 CC The invention comprises a method for delaying the development of, or  
 CC treating a tumour that is associated with human milk fat globules (HMFG)  
 CC or carcinoembryonic antigen (CEA). The method of the invention involves  
 CC administering an anti-idiotypic antibody for HMFG (e.g. 11D10), and an  
 CC anti-idiotypic antibody for CEA (e.g. 3H1). The method is useful for

CC delaying the development, of or treating HMFG/CEA-associated tumours. The  
 CC present cDNA sequence encodes the light chain variable region of the  
 CC mouse 11D10 anti-idiotypic antibody.

XX Sequence 435 BP; 100 A; 111 C; 102 G; 122 T; 0 other;  
 SQ

Query Match 100.0%; Score 435; DB 25; Length 435;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-128;  
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGCCCCCTGCTCAGATTTCTGGGTCTTGTGCTCTGTGTTCCAGTACAGATGT 60  
 DB 1 ATGGGGCCCCCTGCTCAGATTTCTGGGTCTTGTGCTCTGTGTTCCAGTACAGATGT 60  
 |||||  
 QY 61 GACATCCAGATGACCCAGTCTCCATCTCCTTATCTGCCCTCTCTGGGACAAAGAGTCAGT 120  
 |||||  
 DB 61 GACATCCAGATGACCCAGTCTCCATCTCCTTATCTGCCCTCTCTGGGACAAAGAGTCAGT 120  
 |||||  
 QY 121 CTCACCTTGTGGGCAAGTCTAGGACATTTGGTATTAACTTACATTTAGTCTTGGTCTCCCAAA 180  
 |||||  
 DB 121 CTCACCTTGTGGGCAAGTCTAGGACATTTGGTATTAACTTACATTTAGTCTTGGTCTCCCAAA 180  
 |||||  
 QY 181 GATGGAACTATTAAACGCCCTGATCTAGCCACATCCAGTTAGTCTTGGTCTCCCAAA 240  
 |||||  
 DB 181 GATGGAACTATTAAACGCCCTGATCTAGCCACATCCAGTTAGTCTTGGTCTCCCAAA 240  
 |||||  
 QY 241 AGCTTCAGTGGCAGTAGTCTGGTCTCAGATTTATCTCTCACCATCAGCAGCCTTGAGTCT 300  
 |||||  
 DB 241 AGCTTCAGTGGCAGTAGTCTGGTCTCAGATTTATCTCTCACCATCAGCAGCCTTGAGTCT 300  
 |||||  
 QY 301 GAAGATTTGTAGCCTTATCTCTCTACAAATATGCTAGTCTCCGTACAGGTTCCGAGGG 360  
 |||||  
 DB 301 GAAGATTTGTAGCCTTATCTCTCTACAAATATGCTAGTCTCCGTACAGGTTCCGAGGG 360  
 |||||  
 QY 361 GGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
 |||||  
 DB 361 GGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
 |||||  
 QY 421 TCCAGTAAGCTTGGG 435  
 |||||  
 DB 421 TCCAGTAAGCTTGGG 435  
 |||||

## RESULT 4

AAN30165

ID AAN30165 standard; DNA; 450 BP.

XX AAN30165;

XX 25-MAY-1992 (first entry)

XX Sequence encoding the leader, variable region and first 16 AAs of  
 DE the constant region of the kappa-chain (light chain) of MOPC41.

XX Diagnosis; therapy; immunoglobulin; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT CDS 13..78  
 FT /\*tag= a  
 FT /product= leader  
 FT CDS 79..402  
 FT /\*tag= b  
 FT /product= variable region  
 FT 403..450  
 FT /\*tag= c  
 FT /product= constant region

XX EP88994-A.

XX 21-SEP-1983.

XX









```

XX AAT36880;
XX
XX
XX 25-MAR-2003 (updated)
XX 29-OCT-1996 (first entry)
XX
XX 520C9 anti-c-erbB-2 two single chain Fv construct.
XX
XX 520C9; anti-c-erbB-2 monoclonal antibody; single chain Fv; sfv;
XX construct; polypeptide linker; C-terminal amino acid sequence;
XX in vivo imaging; drug targeting experiment; homodimer;
XX increased; binding avidity; tissue retention time; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..732
XX /*tag= a
XX /note= "START codon absent"
XX
XX US5534254-A.
XX
XX 09-JUL-1996.
XX
XX 07-OCT-1993; 93US-0133804.
XX
XX 07-OCT-1993; 93US-0133804.
XX 06-FEB-1992; 92US-0831967.
XX
XX (CHIR ) CHIRON CORP.
XX (CREA-) CREATIVE BIOMOLECULES INC.
XX
XX Houston LL, Huston JS, Oppermann H, Ring DB;
XX
XX WPI: 1996-333194/33.
XX P-PSDB; AAW02280.
XX
XX Compens. contg. antigen-targeting antibody fragment constructs -
XX comprising dimer of single-chain Fv fragments
XX
XX Example 1; Columns 33-36; 30pp; English.
XX
XX Variable heavy (VH) and variable light (VL) genes were cloned from
XX a 520C9 hybridoma cDNA library, using probes directed toward the
XX antibody constant and joining regions. A two single chain Fv (sfv)
XX gene was constructed by connecting the VH and VL genes with a
XX Ser rich polypeptide linker. The resulting 520C9 two sfv gene, the
XX present sequence, was inserted into an expression vector,
XX transformed into E. coli, and protein expression induced by the
XX addn. of IPTG to the culture medium.
XX A compsn. comprising a carrier and the 2 sfv protein prod. can be
XX used for in vivo imaging, and drug targeting experiments. The
XX 2 sfv protein prod. is a homodimer, in which both fragments target
XX the same antigen, therefore giving greater binding avidity and
XX longer tissue retention times, compared to individual sfv protein
XX prod. fragments.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 739 BP; 188 A; 177 C; 181 G; 193 T; 0 other;
XX
Query Match 71.18; Score 309.4; DB 17; Length 739;
Best Local Similarity 95.28; Pred. No. 6.3e-88;
Matches 319; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 56 GATGTGATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGCGACAAAGAG 115
DB 395 GATCCGATATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGCGACAAAGAG 454
QY 116 TACGTCCTCACTGTGCGGGCAAGTCAGGACATTTGTTAATTAACATTTGGCTTCAGCAGG 175
DB 455 TCAGTCTCACTGTGCGGGCAAGTCAGGACATTTGTTAATTAACATTTGGCTTCAGCAGG 514
QY 176 AACGAGATGGAACATATTAAAGCGCTGATCTACGCCACATCCAGTTTGGTCTGGTCTCC 235

```

```

Db 515 AACGAGATGGAACATATTAAACGCGCTGATCTACGCCACATCCAGTTTGGTCTCC 574
QY 236 CCAAAAGGTTTCAGTGGCAGTAGGTCTGGGTTCAGATTATTCTCTCACCATCAGCAGCCTTG 295
Db 575 CCAAAAGGTTTCAGTGGCAGTCGGTCTGGGTTCAGATTATTCTCTCACCATCAGTAGCCTTG 634
QY 296 AGTCTGAAGATTTTGTAGCCCTATTACTGTCTACAAATATGCTAGTCTCCGCTACACGTTTCG 355
Db 635 AGTCTGAAGATTTTGTAGTCTATTACTGTCTACAAATATGCTAGTCTCCGCTACACGTTTCG 694
QY 356 GAGGGGGACCAAGCTGGAATAAAACGGGCTGAT 390
Db 695 GAGGGGGACCAAGCTGGAATAAAACGGGCTGAT 729

RESULT 10
AAV21798
ID AAV21798 standard; cDNA; 739 BP.
XX
XX AAV21798;
XX
XX 16-JUL-1998 (first entry)
XX
XX 520C9 anti-c-erbB-2 sfv' dimeric construct gene sequence.
XX
XX Antigen imaging; single chain Fv; sfv; linker; dimeric; cancer;
XX c-erbB-2; tumour; diagnosis; ss.
XX
XX Synthetic.
XX OS Mus sp.
XX
XX Key Location/Qualifiers
XX CDS 1..732
XX /*tag= a
XX /product= "520C9 sfv' polypeptide"
XX
XX US5753204-A.
XX
XX 19-MAY-1998.
XX
XX 05-JUN-1995; 95US-0461838.
XX
XX 07-OCT-1993; 93US-0133804.
XX 06-FEB-1992; 92US-0831967.
XX 05-JUN-1995; 95US-0461838.
XX
XX (CHIR ) CHIRON CORP.
XX (CREA-) CREATIVE BIOMOLECULES INC.
XX
XX Houston LL, Huston JS, Oppermann H, Ring DB;
XX
XX WPI: 1998-311318/27.
XX P-PSDB; AAW53170.
XX
XX Imaging of antigens in vivo - using dimers of single-chain antibody
XX Fv fragments
XX
XX Example 1; Columns 33-36; 30pp; English.
XX
XX This is the nucleotide sequence of a 520C9 sfv' (single chain Fv)
XX construct. This was constructed by connecting the VH and VL genes with
XX a DNA sequence encoding a Serine rich linker. 520C9 is a monoclonal
XX antibody useful in targeting c-erbB-2 antigen. This dimeric construct
XX can be used in the methods of invention of imaging a preselected antigen
XX expressed in a mammal. The methods are used in magnetic resonance imaging
XX of c-erbB-2 or related antigens in cancer diagnosis. The biosynthetic
XX constructs have enhanced properties as in vivo targeting agents in
XX comparison with intact monoclonal antibodies or their Fab fragments. The
XX dimeric constructs permit the in vivo targeting of an epitope on an
XX antigen with greater apparent avidity, including greater tumour
XX specificity, tumour localisation and same CDRs as the construct.
XX that of the Fab fragment having the same CDRs as the construct.

```

```
XX SQ Sequence 739 BP; 188 A; 177 C; 181 G; 193 T; 0 other;
Query Match 71.1%; Score 309.4; DB 19; Length 739;
Best Local Similarity 95.2%; Pred. No. 6.3e-88;
Matches 319; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 56 GATGTGACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGAG 115
Db 395 GATCCGATATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGAG 454
QY 116 TCAGTCTCACTTGTGGGCAAGTCAGACATTTGGTATTAACTTACATTGGCTTCAGCAGG 175
Db 455 TCAGTCTCACTTGTGGGCAAGTCAGACATTTGGTATTAACTTACATTGGCTTCAGCAGG 514
QY 176 AACCATGATGAACCTATTAAACGCTGATCTACGCCACATCCAGTTTAGTTCTGGTCTCC 235
Db 515 AACCATGATGAACCTATTAAACGCTGATCTACGCCACATCCAGTTTAGTTCTGGTCTCC 574
QY 236 CCAAAAGGTTCAAGTGGCAGTAGTCTCGGGTCAGATTATTCTCTCACCATCAGCAGCCTTG 295
Db 575 CCAAAAGGTTCAAGTGGCAGTAGTCTCGGGTCAGATTATTCTCTCACCATCAGTACGCTTG 634
QY 296 AGTCTGAAGATTTGTAGCTTATTAACAAATAAACGGGCTGAT 390
Db 635 AGTCTGAAGATTTGTAGCTTATTAACAAATAAACGGGCTGAT 729
QY 356 GAGGGGGGACCAAGCTGGAATAAAACGGGCTGAT 390
Db 695 GAGGGGGGACCAAGCTGGAATAAAACGGGCTGAT 729
RESULT 11
AAV63399
ID AAV63399 standard; cDNA; 739 BP.
XX AC AAV63399;
XX DT 28-JAN-1999 (first entry)
XX DE 520C9 sfv DNA sequence.
XX KW 520C9 sfv; antigen; tumour cell; antibody 520C9;
XX KW targeted delivery; antigen-expressing cell; ss.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT CDS 1..732
XX FT /tag= a
XX FT /product= 520C9 sfv
XX PN US5837846-A.
XX PD 17-NOV-1998.
XX PF 05-JUN-1995; 95US-0461386.
XX PR 07-OCT-1993; 93US-0133804.
XX PR 06-FEB-1992; 92US-0831967.
XX PR 05-JUN-1995; 95US-0461386.
XX PA (CHIR ) CHIRON CORP.
XX PA (CREA-) CREATIVE BIOMOLECULES INC.
XX PI Houston LL, Huston JS, Oppermann H, Ring DB;
XX DR WPI; 1999-023541/02.
XX DR P-PSDB; AAW80424.
XX PT Nucleic acid encoding single-chain Fv fragment specific for antigens
XX PT - and having C-terminal tail for crosslinking to form dimer with
XX PT improved pharmacokinetic properties, used to deliver drugs and
```

```
PT imaging agents, especially to tumours
XX Example 1; Columns 33-36; 29pp; English.
XX The present sequence encodes an antibody 520C9 sfv. Variable heavy
CC and light sequences of antibody 7520C9 are connected, together with a
CC serine linker, to produce the present single chain Fv gene. The present
CC sequence exemplifies the invention. Dimers of the single chain Fv are
CC used for targeted delivery of drugs or imaging agents (e.g. cytotoxins,
CC prodrugs or 99m-technetium) to antigen-expressing cells, particularly
CC for treatment or diagnosis of tumours (especially of ovary or breast).
XX SQ Sequence 739 BP; 188 A; 177 C; 181 G; 193 T; 0 other;
Query Match 71.1%; Score 309.4; DB 20; Length 739;
Best Local Similarity 95.2%; Pred. No. 6.3e-88;
Matches 319; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 56 GATGTGACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGAG 115
Db 395 GATCCGATATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGAG 454
QY 116 TCAGTCTCACTTGTGGGCAAGTCAGACATTTGGTATTAACTTACATTGGCTTCAGCAGG 175
Db 455 TCAGTCTCACTTGTGGGCAAGTCAGACATTTGGTATTAACTTACATTGGCTTCAGCAGG 514
QY 176 AACCATGATGAACCTATTAAACGCTGATCTACGCCACATCCAGTTTAGTTCTGGTCTCC 235
Db 515 AACCATGATGAACCTATTAAACGCTGATCTACGCCACATCCAGTTTAGTTCTGGTCTCC 574
QY 236 CCAAAAGGTTCAAGTGGCAGTAGTCTCGGGTCAGATTATTCTCTCACCATCAGCAGCCTTG 295
Db 575 CCAAAAGGTTCAAGTGGCAGTAGTCTCGGGTCAGATTATTCTCTCACCATCAGTACGCTTG 634
QY 296 AGTCTGAAGATTTGTAGCTTATTAACAAATAAACGGGCTGAT 390
Db 635 AGTCTGAAGATTTGTAGCTTATTAACAAATAAACGGGCTGAT 729
QY 356 GAGGGGGGACCAAGCTGGAATAAAACGGGCTGAT 390
Db 695 GAGGGGGGACCAAGCTGGAATAAAACGGGCTGAT 729
RESULT 12
AAZ49543
ID AAZ49543 standard; cDNA; 729 BP.
XX AC AAZ49543;
XX DT 04-APR-2000 (first entry)
XX DE pBscFv#125-2H.HT recombinant cDNA.
XX KW pBscFv#125-2H.HT recombinant cDNA; BscFv#125-2H.HT; interleukin-18;
XX KW immunopathies; inflammatory disorder; autoimmune disease; mouse;
XX KW anti-allergic; anti-inflammatory; immunosuppressive; hematopoietic;
XX KW leukocytopenic; antialgic; antipyretic; ss.
XX OS Mus musculus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT mat_peptide 1..711
XX FT /tag= a
XX FT /note= "single chain variable region fragment
XX FT neutralising IL-18"
XX FT mat_peptide 1..339
XX FT /tag= b
XX FT /note= "Encodes heavy chain variable region"
XX FT mat_peptide 391..711
XX FT /tag= c
XX FT /note= "Encodes light chain variable region"
```



Db 559 GGTGTCCTCCCAAGAGGTTTCAGTGGCAGTAGGTCTGGTCAGATTATCTCCACCATCAGC 618  
Qy 289 AGCTTCAGTCTGAAGATTTTGTAGCCTATTACTGTCTACAATATGCTTCTCCGTAC 348  
Db 619 AGCTTCAGTCTGAAGATTTTGTAGACTATTACTGTCTACAATATGCTTCTCCGTAC 678  
Qy 349 ACCTTCGGAGGGGGACCAAGCTGGAATAAAA 381  
Db 679 ACCTTCGGAGGGGGACCAAGCTGGAATAAAA 711

## RESULT 14

AAZ49534  
ID AAZ49534 standard; cDNA: 324 BP.

XX AC AAZ49534;

XX DT 04-APR-2000 (first entry)

XX DE Mouse anti-IL-18 antibody VL encoding cDNA from hybridoma #125-2H.

XX KW Anti-interleukin-18 antibody; IL-18; light chain variable region; VL;  
KW hybridoma #125-2H; mouse; monoclonal antibody #125-2HmAb; interleukin-18;  
KW antiinflammatory; immunosuppressive; leucocytopenic; antialgic;  
KW antipyretic; antiallergic; hepatotropic; treatment; autoimmune disease;  
KW immunopathy; inflammatory disorder; immunoreaction; ss.

XX OS Mus musculus.

XX FH Key Location/Qualifiers

FT mat\_peptide 1..324

FT /\*tag= a

FT /label= Anti-IL-18\_antibody\_light\_chain\_variable\_region

XX PN EP974600-A2.

XX PD 26-JAN-2000.

XX PF 24-JUN-1999; 99EP-0304977.

XX PR 24-JUN-1998; 98JP-0177580.

XX PR 12-OCT-1998; 98JP-0289044.

XX PR 22-DEC-1998; 98JP-0365023.

XX PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.

XX PI Nishida Y, Okura T, Tanimoto T, Kurimoto M;

XX DR WPI; 2000-118341/11.

XX DR P-PSDB; AAY44587.

XX PT New artificially produced peptide for neutralizing biological activity  
PT of interleukin-18, useful for treating and preventing immunopathies,  
PT inflammatory disorders and autoimmune diseases -

XX PS Claim 11; Page 21; 36pp; English.

XX CC The present cDNA sequence derived from hybridoma #125-2H, encodes mouse  
CC anti-interleukin-18 antibody light chain variable region (VL). It can be  
CC used in the production of recombinant monoclonal antibody #125-2HmAb,  
CC which is capable of neutralising biological activities of interleukin-18.  
CC The antibody has antinflammatory, immunosuppressive, leucocytopenic,  
CC antialgic, antipyretic, antiallergic and hepatotropic activity and can be  
CC used for prevention and treatment of autoimmune diseases, immunopathies  
CC and inflammatory disorders caused by excessive immunoreaction.

XX SQ Sequence 324 BP; 83 A; 78 C; 76 G; 87 T; 0 other;

Query Match

Best Local Similarity 70.4%; Score 306.4; DB 21; Length 324;

Matches 313; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 61 GACATCCAGATGACCCAGTCTCCATCTCTCTATCTCTCTGGACAAAGAGTCAGT 120  
Db 1 GACATCCAGATGACCCAGTCTCCATCTCTCTATCTCTCTGGAGAAAGAGTCAGT 60  
Qy 121 CTCACCTTGTGCGGCAAGTCAGGACATTGGTATTAACTTACATTGGCTTCAGCAGAACCA 180  
Db 61 CTCACCTTGTGCGGCAAGTCAGGACATTGGTATTAACTTACATTGGCTTCAGCAGAACCA 120  
Qy 181 GATGGAACTATTAAACGCCTGATCTACGCCACATCCAGTTTAGCTTCTGGTGTCCCAAAA 240  
Db 121 GATGGAACTTTTAAACGCCCTGATCTACGCCACATCCAGTTTAGCTTCTGGTGTCCCAAG 180  
Qy 241 AGTTTCAGTGGCAGTAGTCTGGCTCAGATTATTTCTTCACCATCAGCAGCCTTGTAGTCT 300  
Db 181 AGTTTCAGTGGCAGTAGTCTGGCTCAGATTATTTCTTCACCATCAGCAGCCTTGTAGTCT 240  
Qy 301 GAAGATTTTGTAGCCTATTACTGTCTACAATATGCTAGTTCTCGGTACACGTTCCGGAGGG 360  
Db 241 GAAGATTTTGTAGCCTATTACTGTCTACAATATGCTAGTTCTCGGTACACGTTCCGGAGGG 300  
Qy 361 GGGACCAAGCTGGAAATAAAAACGG 384  
Db 301 GGGACCAAGCTGGCAATAAAAACGG 324

## RESULT 15

AAQ46086  
ID AAQ46086 standard; cDNA: 1605 BP.

XX AC AAQ46086;

XX DT 25-MAR-2003 (updated)

XX DT 07-FEB-1994 (first entry)

XX DE Sequence encoding G-FIT.

XX KW Tumour antigen; c-erbB-2; G-FIT; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT CDS 1..1605

FT /\*tag= a

XX PN WO9316185-A2.

XX PD 19-AUG-1993.

XX PF 05-FEB-1993; 93WO-US01055.

XX PR 06-FEB-1992; 92US-0831967.

XX PA (CETU ) CETUS ONCOLOGY CORP.

XX PA (CREA-) CREATIVE BIOMOLECULES INC.

XX PI Houston LL, Huston JS, Oppermann H, Ring DB;

XX DR WPI; 1993-272889/34.

XX DR P-PSDB; AAR39571.

XX PT New single chain Fv polypeptide binding to C-erbB-2 tumour  
PT antigen - for imaging or treating breast or ovarian cancer etc.

XX PS Example; pages 65-68; 87pp; English.

XX CC c-erbB-2 refers to a protein antigen expressed on the surface of  
CC tumour cells, such as breast and ovarian tumour cells, which is an  
CC approx. 200,000 mol. wt. acidic glycoprotein having an isoelectric  
CC pt. of about 5.3 (see AAQ46083, AAR39568).

XX CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 1605 BP; 434 A; 354 C; 373 G; 444 T; 0 other;

Query Match		69.4%;	Score 302;	DB 14;	Length 1605;
Best Local Similarity		92.7%;	Pred. No. 1.9e-85;		
Matches 317;		Conservative	0;	Mismatches 25;	Indels 0; Gaps 0;
Qy	49	GGTACCAGATGTGACATCCAGATGACCCAGTCTCCATCCTTATCTGCTCTCTGCGGA	108		
Db	1261	GGCGGGGATCGGATATCCAGATGACCCAGTCTCCATCCTCTTATCTGCTCTCTGCGGA	1320		
Qy	109	CAAAGAGTCACTCTCACTTCTCGGCAAGTCAGGACATTTGGTATTAACTTACATTGGCTT	168		
Db	1321	GAAGAGTCACTCTCACTTCTCGGCAAGTCAGGACATTTGGTATTAACTTACATTGGCTT	1380		
Qy	169	CAGCAGAACCCAGATGGAATATTAAACGCTGATCTACGCCACATCCAGTTTAGGTTCT	228		
Db	1381	TCACAGGAACAGATGGAATATTAAACGCTGATCTACGCCACATCCACTTTAGATTCT	1440		
Qy	229	GGTGTCGCCAAAGGTTCACTGGCAGTAGTCTGGGTCAGATTATTCTCACCACAGC	288		
Db	1441	GGTGTCGCCAAAGGTTCACTGGCAGTAGTCTGGGTCAGATTATTCTCACCACAGT	1500		
Qy	289	AGCCTTGAGTCTGAAGATTTTGTAGCCTATTACTGTCTACAATATGCTAGTTCTCCGTAC	348		
Db	1501	AGCCTTGAGTCTGAAGATTTTGTAGTCTATTACTGTCTACAATATCCTATTTTCCGTAC	1560		
Qy	349	ACGTTGGAGGGGGACCAAGCTGGAATAAAACGGGCTGAT	390		
Db	1561	ACGTTGGAGGGGGACCAACCTGGAAATAAAACGGGCTGAT	1602		

Search completed: August 30, 2003, 18:52:34  
Job time : 192.798 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 17:29:10 ; Search time 159,241 Seconds  
(without alignments)  
6267.222 Million cell updates/sec

Title: US-08-836-455-1  
Perfect score: 435  
Sequence: 1 ATGGGGGCCCTGCTCAGAT.....CACCATCCAGTAAGCTTGGG 435

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09D\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	435	100.0	435	9	US-09-861-294-1
2	435	100.0	435	12	US-10-367-506-1
3	384.6	88.4	407	10	US-09-924-099-27
4	372	85.5	402	8	US-08-779-784-5
5	329.6	75.8	390	12	US-10-268-883-10
6	323.4	74.3	387	9	US-09-007-093-1
7	323.2	74.3	351	8	US-08-779-784-17
8	322.6	74.2	351	8	US-08-779-784-16
9	309.4	71.1	739	10	US-09-887-853-5
10	308.4	70.9	729	10	US-09-924-099-20
11	307.4	70.7	711	10	US-09-924-099-19
12	306.4	70.4	324	9	US-09-924-099-11
13	267.6	61.5	447	9	US-09-797-481-1
14	267.6	61.5	447	14	US-09-844-736-3
15	267.6	61.5	447	14	US-10-162-396-3
16	266	61.1	447	9	US-09-861-294-19

17	266	61.1	447	12	US-10-367-506-19	Sequence 19, Appl
18	265.2	61.0	490	11	US-09-918-995-37859	Sequence 37859, A
19	263.6	60.6	714	14	US-10-153-382-18	Sequence 18, Appl
20	261.6	60.1	520	10	US-09-878-178-1210	Sequence 1210, Ap
21	261.6	60.1	520	13	US-10-046-935-1210	Sequence 1210, Ap
22	261.6	60.1	520	14	US-10-146-935-1210	Sequence 1210, Ap
23	258.8	59.5	514	14	US-10-066-543-2025	Sequence 2025, Ap
24	258.8	59.5	537	14	US-10-066-543-186	Sequence 186, App
25	258.8	59.5	974	10	US-09-859-053-29	Sequence 29, Appl
26	258.6	59.4	393	8	US-08-779-784-20	Sequence 20, Appl
27	257.2	59.1	729	14	US-10-216-484-125	Sequence 125, App
28	252.2	58.0	737	10	US-09-919-344-7	Sequence 7, Appl
29	250.6	57.6	321	14	US-10-056-794-16	Sequence 16, Appl
30	249.6	57.4	634	14	US-10-158-646-55	Sequence 55, Appl
31	248	57.0	408	9	US-09-764-304-9	Sequence 9, Appl
32	248	57.0	408	14	US-10-265-713-9	Sequence 9, Appl
33	247.6	56.9	705	9	US-09-740-002-16	Sequence 16, Appl
34	246.4	56.6	408	14	US-10-193-752-19	GENERAL INFORMA
35	246	56.6	536	14	US-10-198-846-8024	Sequence 8024, Ap
36	246	56.6	698	10	US-09-844-684-11	Sequence 11, Appl
37	246	56.6	698	14	US-10-040-244-11	Sequence 11, Appl
38	246	56.6	728	10	US-09-844-684-15	Sequence 15, Appl
39	246	56.6	728	14	US-10-040-244-15	Sequence 15, Appl
40	246	56.6	819	14	US-10-158-646-65	Sequence 65, Appl
41	244.4	56.2	716	10	US-09-844-684-13	Sequence 13, Appl
42	244.4	56.2	716	14	US-10-040-244-13	Sequence 13, Appl
43	244	56.1	990	9	US-09-800-729-79	Sequence 79, Appl
44	243.4	56.0	381	12	US-10-268-883-4	Sequence 4, Appl
45	241.8	55.6	381	14	US-10-084-139-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-861-294-1  
; Sequence 1, Application US/09861294  
; Patent No. US20020098190A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; FILE OF INVENTION: TUMORS BEARING HMFG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/09/861,294  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 435  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(435)  
; NAME/KEY: sig\_peptide  
; LOCATION: (1)...(60)  
; NAME/KEY: mat\_peptide  
; LOCATION: (61)...(435)  
US-09-861-294-1

Query Match 100.0%; Score 435; DB 9; Length 435;  
Best Local Similarity 100.0%; Pred. No. 2.7e-136;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGGGCCCCGCTCAGATCTTGGTCTTGTGCTTGTTCAGGTACCATGT 60  
|||||  
DB 1 ATGGGGCCCCGCTCAGATCTTGGTCTTGTGCTTGTTCAGGTACCATGT 60  
|||||

Qy	61	GACATCCAGATGACCCAGCTCCCATCCTCCTTATCTGCCTCTCTGGGACAAAGAGTCAGT	120
Db	61	GACATCCAGATGACCCAGCTCCCATCCTCCTTATCTGCCTCTCTGGGACAAAGAGTCAGT	120
Qy	121	CTCACCTTCTGGGCAAGTCAGGACATTTGGTATTAACTTACATTGGCTTCAGCAGGACCA	180
Db	121	CTCACCTTCTGGGCAAGTCAGGACATTTGGTATTAACTTACATTGGCTTCAGCAGGACCA	180
Qy	181	GATGGAACCTATTAAACGGCTGATCTACGCCACATCCAGTTTAGTTTCTGGTGTCCCCAAA	240
Db	181	GATGGAACCTATTAAACGGCTGATCTACGCCACATCCAGTTTAGTTTCTGGTGTCCCCAAA	240
Qy	241	AGGTTTCAGTGGCAGTAGTCTGGGTCAAGATTATCTCTACCATCAGCAGCCTTGAGTCT	300
Db	241	AGGTTTCAGTGGCAGTAGTCTGGGTCAAGATTATCTCTACCATCAGCAGCCTTGAGTCT	300
Qy	301	GAAGATTTTGTAGCCTATTACTGTCTACATATGCTAGTTCTCCGTACAGTTCGGAGGG	360
Db	301	GAAGATTTTGTAGCCTATTACTGTCTACATATGCTAGTTCTCCGTACAGTTCGGAGGG	360
Qy	361	GGGACCAAGCTGGAAATAAAACGGGCTGATCTGCACCAACTGTATCCATCTTCCCACCA	420
Db	361	GGGACCAAGCTGGAAATAAAACGGGCTGATCTGCACCAACTGTATCCATCTTCCCACCA	420
Qy	421	TCCAGTAAGCTTTGGG	435
Db	421	TCCAGTAAGCTTTGGG	435

```

RESULT 2
US-10-367-506-1
; Sequence 1, Application US/10367506
; Publication NO. US20030152575A1
; GENERAL INFORMATION:
; APPLICANT: Malaya CHATTERJEE
; APPLICANT: Kenneth A. FOON
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: TUMORS BEARING HMGF AND CEA ANTIGENS
; FILE REFERENCE: 304142000620
; CURRENT APPLICATION NUMBER: US/10/367,506
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US/09/861,294
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/049,540
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 09/096,244
; PRIOR FILING DATE: 1998-06-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(435)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (1)...(60)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (61)...(435)
US-10-367-506-1

```

	Query Match	100.0%	Score 435;	DB 12;	Length 435;
	Best Local Similarity	100.0%;	Pred. No. 2.7e-136;		
	Matches 435;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGGGGGCCCTGCTCAGATTC	TGGGGTCTCTGTGCTGTGTTCCAGGTACCA	GATG	60
Db	1	ATGGGGGGCCCTGCTCAGATTC	TGGGGTCTCTGTGCTGTGTTCCAGGTACCA	GATG	60
QY	61	GACATCCAGATGACCCAGCTCTCCATCTCTCTATCTGCCCTCTCTGGGACAAAGAGT	CACT	120	

Db	61		GACATCCAGATGACCAGCTCCATCCCTCCCTATCTGCCCTCTGGGACAAAGAGTCAGT	120
Qy	121		CTCAGCTTGTGGGCAAGTCAGGACATTGGTATTAACTTACATTTAGCTTTCAGCAGGAACCA	180
Db	121		CTCAGCTTGTGGGCAAGTCAGGACATTGGTATTAACTTACATTTAGCTTTCAGCAGGAACCA	180
Qy	181		GATGGAACCTATTAAACGGCTGATCTACGCCACATCCAGTTTAGTTCTCGGTGTCGCCCAA	240
Db	181		GATGGAACCTATTAAACGGCTGATCTACGCCACATCCAGTTTAGTTCTCGGTGTCGCCCAA	240
Qy	241		AGGTTTCAGTGGCAGTAGTCTGGGTCAGATTATTCTCTCACCATCAGCAGCCTTTGAGTCT	300
Db	241		AGGTTTCAGTGGCAGTAGTCTGGGTCAGATTATTCTCTCACCATCAGCAGCCTTTGAGTCT	300
Qy	301		GAAGATTTTGTAGCCTATTACTGTCTCAAAATATGCTAGTTCTCCGTACAGTTTCGGAGGG	360
Db	301		GAAGATTTTGTAGCCTATTACTGTCTCAAAATATGCTAGTTCTCCGTACAGTTTCGGAGGG	360
Qy	361		GGGACCAAGCTGGAAATAAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA	420
Db	361		GGGACCAAGCTGGAAATAAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA	420
Qy	421		TCCAGTAAGCTTGGG	435
Db	421		TCCAGTAAGCTTGGG	435

RESULT 3  
US-09-924-099-27  
; Sequence 27, Application US/09924099  
; Patent No. US20020128450A1  
; GENERAL INFORMATION:  
; APPLICANT: NISHIDA, Yoshihiro  
; APPLICANT: OKURA, Takanori  
; APPLICANT: TANIMOTO, Tadao  
; APPLICANT: KORIMOTO, Masashi  
; TITLE OF INVENTION: PEPTIDE  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/924,099  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 33  
; SEQ. ID. NO. 27

```

? LENGTH: 407
?
? TYPE: DNA
? ORGANISM: Mus musculus
?
? FEATURE:
?
? NAME/KEY: CDS
?
? LOCATION: (1)...(407)
? NAME/KEY: sig peptide
? LOCATION: (1)...(60)
?
? US-09-924-099-27

```

Query Match 88.4%; Score 384.6; DB 10; Length 407;  
Best Local Similarity 96.6%; Pred. No. 2.6e-119;  
Matches 393; Conservative 0; Mismatches 14; Indels 0; Gaps 0

Qy	1	ATGGGGCCCCGCTGCAGATCTTGGGTTCTTGTCTTGTTCAGGTACAGATGT	60
Db	1	ATGAGGCCCTGCCTCAGATTTTGGCTCTTGTGCTGTGTTCCAGGTACAGATGT	60
Qy	61	GACATCCAGATGACCCAGTCTCCATCTCCTTATCTGCCTCTCTGGGACAAAGATCAGT	120
Db	61	GACATCCAGATGACCCAGTCTCCATCTCCTTATCTGCCTCTCTGGGAAAGATCAGT	120
Qy	121	CTCAGTTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTGGCTTCAGCAGGAACCA	180



Db 121 CTCACCTGTCGGCAAGTCAGGACATTTGGTAGTAAATATATACATGGCTTCAACAGGAACCA 180  
QY 181 GATGGAACTATTAAACGCCCTGATCTAGCCACATCCAGTTTAGTGTCTGGTGTCCCAAA 240  
Db 181 GATGGAACTATTAAACGCCCTGATCTAGCCACATCCAGTTTAGTGTCTGGTGTCCCAAG 240  
QY 241 AGGTTGAGTGGCAGTAGGTCCTGAGATTTATCTCTACCATCAGCAGCCTTGAGTCT 300  
Db 241 AGGTTGAGTGGCAGTAGGTCCTGAGATTTATCTCTACCATCAGCAGCCTTGAGTCT 300  
QY 301 GAAGATTTGAGCCTTACTCTTACAAATAGCTAGTCTCCGTACAGGTCGAGGG 360  
Db 301 GAAGATTTGAGCCTTACTCTTACAAATAGCTAGTCTCCGTACAGGTCGAGGG 360  
QY 361 GGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACTGATC 407  
Db 361 GGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACTGATC 407

## RESULT 4

US-08-779-784-5  
; Sequence 5, Application US/08779784  
; Publication No. US20020164325A1  
; GENERAL INFORMATION:  
; APPLICANT: Rodriguez, Moses  
; APPLICANT: Miller, David J.  
; APPLICANT: Asakura, Kunihiko  
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM  
; TITLE OF INVENTION: REMYELINATION USING MONOCLONAL AUTOANTIBODIES  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/779,784  
; FILING DATE: 07-JAN-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/692,084  
; FILING DATE: 08-AUG-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/236,520  
; FILING DATE: 29-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 402 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEetical: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Mus musculus  
US-08-779-784-5

Query Match

85.5%; Score 372; DB 8; Length 402;

Best Local Similarity 96.2%; Pred. No. 4.6e-115;  
Matches 381; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 1 ATGGGGCCCCCTGCTCAGATTTCTGGTTCCTTCTGCTGCTTCTTCCAGGTACAGATGT 60  
Db 7 ATAGGGGCTCTCTGACAGATTTTGGCTTCTTCTGCTGCTTCTTCAAGGTACAGATGT 66  
QY 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGAGTCAGT 120  
Db 67 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGAGAAAGAGTCAGT 126  
QY 121 CTCACCTGTCGGCAAGTCAGGACATTTGGTAGTAAATATATACATGGCTTCAACAGGAACCA 180  
Db 127 CTCACCTGTCGGCAAGTCAGGACATTTGGTAGTAAATATATACATGGCTTCAACAGGAACCA 186  
QY 181 GATGGAACTATTAAACGCCCTGATCTAGCCACATCCAGTTTAGTGTCTGGTGTCCCAAA 240  
Db 187 GATGGAACTATTAAACGCCCTGATCTAGCCACATCCAGTTTAGTGTCTGGTGTCCCAAA 246  
QY 241 AGGTTGAGTGGCAGTAGGTCCTGAGATTTATCTCTACCATCAGCAGCCTTGAGTCT 300  
Db 247 AGGTTGAGTGGCAGTAGGTCCTGAGATTTATCTCTACCATCAGCAGCCTTGAGTCT 306  
QY 301 GAAGATTTGAGCCTTACTCTTACAAATAGCTAGTCTCCGTACAGGTCGAGGG 360  
Db 307 GAAGATTTGAGCCTTACTCTTACAAATAGCTAGTCTCCGTACAGGTCGAGGG 366  
QY 361 GGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACTGATC 396  
Db 367 GGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACTGATC 402

## RESULT 5

US-10-268-883-10  
; Sequence 10, Application US/10268883  
; Publication No. US20030138862A1  
; GENERAL INFORMATION:  
; APPLICANT: Tso, J. Yun  
; APPLICANT: Green, Jennifer Macphate  
; TITLE OF INVENTION: Anti-HLA-DR Antibodies and the Methods of Using Thereof  
; FILE REFERENCE: 05882,0062.NPUS01  
; CURRENT APPLICATION NUMBER: US/10/268,883  
; CURRENT FILING DATE: 2003-03-26  
; PRIOR APPLICATION NUMBER: USSN 60/329,178  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: USSN 60/331,965  
; PRIOR FILING DATE: 2001-11-21  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 10  
; LENGTH: 390  
; TYPE: DNA  
; ORGANISM: Mouse  
US-10-268-883-10

Query Match. 75.8%; Score 329.6; DB 12; Length 390;  
Best Local Similarity 91.1%; Pred. No. 9e-101;  
Matches 350; Conservative 0; Mismatches 34; Indels 0; Gaps 0;  
QY 1 ATGGGGCCCCCTGCTCAGATTTCTGGTTCCTTCTGCTGCTTCTTCCAGGTACAGATGT 60  
Db 7 ATAGGGGTTCTCTGCTCAGGTTTGGCTTCTTCTGCTGCTTCTTCCAGGTACAGATGT 66  
QY 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGAGTCAGT 120  
Db 67 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGAGAAAGAGTCAGT 126  
QY 121 CTCACCTGTCGGCAAGTCAGGACATTTGGTAGTAAATATATACATGGCTTCAACAGGAACCA 180  
Db 127 CTCACCTGTCGGCAAGTCAGGAAATTTAGTGGTTACTTAAGTGGCTTCAACAGGAACCA 186  
QY 181 GATGGAACTATTAAACGCCCTGATCTAGCCACATCCAGTTTAGTGTCTGGTGTCCCAAA 240

Db 187 GATGGAACATATTAACCGCTGATCTACGCCCATCCACTTTAGATTCTGTGTCCTCCAAAA 246  
Qy 241 AGGTTTCAGTCAGTAGTCTGGCTCAGATTATCTCTCACCATCAGCAGCTTGGAGTCT 300  
Db 247 AGGTTTCAGTCAGTAGTCTGGCTCAGATTATCTCTCACCATCAGCAGCTTGGAGTCT 306  
Qy 301 GAAGATTTTGTAGCTTATCTGCTCAATATCTGCTTCCGTACACGCTTCCGAGGG 360  
Db 307 GAAGATTTTGCAGACTATCTGCTCAATATCTGCTTCCGTACACGCTTCCGAGGG 366  
Qy 361 GGGACCAAGCTGGAATAAAGCG 384  
Db 367 GGCACCAAGCTGGAATAAAGCG 390

## RESULT 6

US-09-007-093-1  
; Sequence 1, Application US/09007093  
; Patent No. US2002025315A1  
; GENERAL INFORMATION:  
; APPLICANT: Anand, Naveen N  
; APPLICANT: Barber, Brian H  
; APPLICANT: Cates, George A  
; APPLICANT: Caterini, Judith E  
; APPLICANT: Klein, Michel H  
; TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF  
; TITLE OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/007,093  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION NUMBER: US 08/483,576  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-765  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 387 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-007-093-1

Query Match 74.3%; Score 323.4; DB 9; Length 387;  
Best Local Similarity 90.6%; Pred. No. 1.1e-96;  
Matches 345; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 1 ATGGGGCCCTGCTCAGATCTGGCTTCTTGTCTCTTGTTCAGGTACAGATGT 60  
Db 7 ATGAGGGTCTCTGCTCAGTTTGTGGCTTCTTGTCTCTTGTTCAGGTACAGATGT 66  
Qy 61 GACATCCAGATGACCCAGTCTCCATCTCTATCTGCTCTCTGGGACAAAGAGTCA 120  
Db 67 GACATCCAGATGACCCAGTCTCCATCTCTATCTGCTCTCTGGGACAAAGAGTCA 126

Qy 121 CTCACTTCTCGGCAAGTCAGGACATTTGGTATTAACTTACATTGCTTTCAGCAGAACCA 180  
Db 127 CTCATTGTCTCGGCAAGTCAGGAAATAGTGGTTACTTAACTTTCAGCAGAACCA 186  
Qy 181 GATGGAACATATTAACCGCTGATCTACGCCCATCCAGTTTAGTTCGTGGTGTCCCAAA 240  
Db 187 GATGGAACATATTAACCGCTGATCTACGCCCATCCAGTTTAGTTCGTGGTGTCCCAAA 246  
Qy 241 AGGTTTCAGTCAGTAGTCTGGCTCAGATTATCTCTCACCATCAGCAGCTTGGAGTCT 300  
Db 247 AGGTTTCAGTCAGTAGTCTGGCTCAGATTATCTCTCACCATCAGCAGCTTGGAGTCT 306  
Qy 301 GAAGATTTTGTAGCTTATCTGCTCAATATCTGCTTCCGTACACGCTTCCGAGGG 360  
Db 307 GAAGATTTTGCAGACTATCTGCTCAATATCTGCTTCCGTACACGCTTCCGAGGG 366  
Qy 361 GGGACCAAGCTGGAATAAATAA 381  
Db 367 GGGACCAAGCTGGAGCTGAA 387

## RESULT 7

US-08-779-784-17  
; Sequence 17, Application US/08779784  
; Publication No. US20020164325A1  
; GENERAL INFORMATION:  
; APPLICANT: Rodriguez, Moses  
; APPLICANT: Miller, David J.  
; APPLICANT: Asakura, Kunihiro  
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM  
; TITLE OF INVENTION: REMYELINATION USING MONOCLONAL AUTOANTIBODIES  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; STREET: Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/779,784  
; FILING DATE: 07-JAN-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/692,084  
; FILING DATE: 08-AUG-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/236,520  
; FILING DATE: 29-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 351 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; HYPOTHETICAL: NO  
US-08-779-784-17









Search completed: August 30, 2003, 19:50:20  
Job time : 161.241 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 16:08:05 ; Search time 1512.79 seconds  
(without alignments)  
6988.707 Million cell updates/sec

Title: US-08-836-455-1  
Perfect score: 435  
Sequence: 1 ATGGGGGCCCTGCTCAGAT.....CACCATCCAGTAAGCTTGGG 435

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_esti:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	373.2	85.8	630	10	BF138788
2	362.8	83.4	926	12	BG962572
3	334	76.8	714	12	BG963548
4	320.2	73.6	374	13	BY083003

5	287.8	66.2	827	12	BI152061
6	287.6	66.1	874	10	BG518527
7	286.8	65.9	891	10	BF579422
8	285.2	65.6	626	10	BF582283
9	285.2	65.6	725	10	BF580940
10	282.8	65.0	1459	13	BQ918407
11	278.8	64.1	762	10	BF144806
12	278	63.9	901	13	BQ943305
13	276.4	63.5	772	10	BF581992
14	276.4	63.5	830	12	BI455041
15	275.6	63.4	695	10	BE284224
16	274.8	63.2	669	10	BF123422
17	274.8	63.2	708	12	BI103114
18	273.2	62.8	886	10	BG756818
19	268.8	61.8	685	10	BG518543
20	266.8	61.3	989	13	BUR99279
21	265.2	61.0	484	2	HSW070437
22	265.2	61.0	764	14	CB956251
23	265.2	61.0	1019	12	BM914405
24	264.4	60.8	354	13	BY085718
25	263.6	60.6	723	14	CB956143
26	263.6	60.6	922	10	BF976253
27	262.8	60.4	851	12	BI101548
28	262.8	60.4	1038	10	BG757218
29	262.2	60.3	966	10	BF578083
30	262	60.2	837	14	CB984807
31	262	60.2	867	10	BG754732
32	262	60.2	939	13	BO705876
33	260.6	59.9	691	10	BE307937
34	260.4	59.9	657	14	CB955665
35	260.4	59.9	763	14	CB985168
36	260.4	59.9	770	14	CB987520
37	260.4	59.9	776	14	CB986151
38	260.4	59.9	805	14	CB955618
39	260.4	59.9	807	14	CB958795
40	260.4	59.9	888	13	BQ881840
41	260.4	59.9	920	13	BUR99295
42	258.8	59.5	472	9	AW406323
43	258.8	59.5	493	9	AW405753
44	258.8	59.5	599	10	BE287315
45	258.8	59.5	724	14	CB959008

ALIGNMENTS

RESULT 1  
BF138788  
LOCUS  
DEFINITION 601780387F1 NCI\_CGAP\_Lu30 Mus musculus cDNA clone IMAGE:4008404 5',  
mRNA sequence.  
ACCESSION BF138788  
VERSION BF138788.1 GI:10977828  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 630)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LIAW9242 row: n column: 21  
High quality sequence stop: 628.

BF138788 630 bp mRNA linear EST 24-OCT-2000  
601780387F1 NCI\_CGAP\_Lu30 Mus musculus cDNA clone IMAGE:4008404 5',  
mRNA sequence.

```

FEATURES          Location/Qualifiers
source
1..630
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="CZECH II"
  /db_xref="taxon:10090"
  /clone="IMAGE:4008404"
  /tissue_type="tumor, metastatic to mammary"
  /lab_host="DHI0B"
  /clone_lib="NCI_CGAP_Lu30"
  /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
  Site_2: SalI; transgenic model WNT-1, expression driven by
  MMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
  dt. Library constructed by Life Technologies.
  Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT      162 a 158 c 150 g 160 t
ORIGIN
Query Match      85.8%; Score 373.2; DB 10; Length 630;
Best Local Similarity 92.3%; Pred. NO. 4.3e-106;
Matches 393; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 5 GGGCCCTGCTCAGATCTTGGGTCTTGTGCTCTGTTCCAGGTACCAAGATGTGACA 64
Db 1 GGGCTCTCGCCAGGTTTTGGCTTCTTGTGCTCTGTTCCAGGTGCCAGATGTGACA 60
QY 65 TCCAGATGACCCAGTCTCCATCCCTTATCTGCTCTCTCTGGGACAAAGATCACTCA 124
Db 61 TCCAGATGACCCAGTCTCCATCCCTTATCTGCTCTCTCTGGGAGAAAGATCACTCA 120
QY 125 CTTGTGGGCAAGTCAGACATTTGGTATTAATTAATTAATTAATTAATTAATTAAT 184
Db 121 CTTGTGGGCAAGTCAGACATTTGGTATTAATTAATTAATTAATTAATTAATTAAT 180
QY 185 GAATATTAAAGCGCTGATCTACGCCACATCCAGTTTGGTCTGTTCCCAAAAGGT 244
Db 181 GAATATTAAAGCGCTGATCTACGCCACATCCAGTTTGGTCTGTTCCCAAAAGGT 240
QY 245 TCAGTGGCAGTAGTCTGGGTGAGATTATTTCTCACCATCAGCAGCCTTGAGTCTGAAG 304
Db 241 TCAGTGGCAGTAGTCTGGGTGAGATTATTTCTCACCATCAGCAGCCTTGAGTCTGAAG 300
QY 305 ATTTGTAGCTATTAATGCTTACAATGATGATGCTGACCAACTGTATCCATCTTCCACCATCCA 424
Db 301 ACTTGTGACAGATTAATGCTTACAATGATGATGCTGACCAACTGTATCCATCTTCCACCATCCA 420
QY 425 GTAAGC 430
Db 421 GTGAGC 426

RESULT 2
BG962572          926 bp  mRNA  linear  EST 12-JUN-2001
LOCUS             60282953F1 NCI_CGAP_Co24 Mus musculus cdna clone IMAGE:4984788 5',
DEFINITION        mRNA sequence.
ACCESSION         BG962572
VERSION           BG962572.1 GI:14350209
KEYWORDS          EST.
SOURCE            Mus musculus (house mouse)
ORGANISM          Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE         1 (bases 1 to 926)
AUTHORS          NIH-MGC http://mgc.nci.nih.gov/.
TITLE            National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL          Unpublished
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10991 row: 1 column: 13
High quality sequence stop: 685.

FEATURES          Location/Qualifiers
source
1..926
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="FVB/N"
  /db_xref="taxon:10090"
  /clone="IMAGE:4984788"
  /lab_host="DHI0B (T1 phage-resistant)"
  /clone_lib="NCI_CGAP_Co24"
  /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
  Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
  Average insert size 1.6 kb. Constructed by Life
  Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      230 a 259 c 205 g 232 t
ORIGIN
Query Match      83.4%; Score 362.8; DB 12; Length 926;
Best Local Similarity 90.2%; Pred. NO. 9.4e-103;
Matches 388; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 1 ATGGGGGCCCTGCTCAGATCTTGGGTCTTGTGCTCTGTTCCAGGTACCAAGATGT 60
Db 15 ATGAGGGTCTCTCTCAGTCTTGGGTCTTGTGCTCTGTTCCAGGTACCAAGATGT 74
QY 61 GACATCCAGATGACCCAGTCTCCATCCCTTATCTGCTCTCTCTGGGACAAAGATCACT 120
Db 75 GACATCCAGATGACCCAGTCTCCATCCCTTATCTGCTCTCTCTGGGAGAAAGATCACT 134
QY 121 CTCATCTGTTCGGGCAAGTCAGGACATTTGGTATTAATTAATTAATTAATTAATTAAT 180
Db 135 CTCATCTGTTCGGGCAAGTCAGGAGTTGGTCTTCTTAAAGTTGGCTTCAGCAAAACCA 194
QY 181 GATGAACATATAAAGCGCTGATCTACGCCACATCCAGTTTGGTCTGTTCCCAAA 240
Db 195 GATGAACATTTAAAGCGCTGATCTACGCCACATCCAGTTTGGTCTGTTCCCAAA 254
QY 241 AGTTTCAGTGGCAGTAGTCTGGGTGAGATTATTTCTCACCATCAGCAGCCTTGAGTCT 300
Db 255 AGTTTCAGTGGCAGTAGTCTGGGTGAGATTATTTCTCACCGTCAGTTCAGTTCAGTCT 314
QY 301 GAAGATTTGTAGCTTATTAATGCTTACAATGATGATGCTTCCGTACACGTTCCGAGGG 360
Db 315 GAAGATTTGTAGCTTATTAATGCTTACAATGATGATGCTTCCGTACACGTTCCGAGGG 374
QY 361 GGGACCAAGCTGGAATAAAACGGGCTGATCTGACCAACTGTATCCATCTTCCACCA 420
Db 375 GGCACAGGCTGGAATCAAAACGGGCTGATCTGACCAACTGTATCCATCTTCCACCA 434
QY 421 TCCAGTAAGC 430
Db 435 TCCAGTGAGC 444

RESULT 3
BG963548          714 bp  mRNA  linear  EST 12-JUN-2001
LOCUS             602831226F1 NCI_CGAP_Co24 Mus musculus cdna clone IMAGE:4985791 5',
DEFINITION        mRNA sequence.
ACCESSION         BG963548
VERSION           BG963548.1 GI:14351185
KEYWORDS          EST.
SOURCE            Mus musculus (house mouse)
ORGANISM          Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	NIH-MGC	NIH-MGC	
1	(bases 1 to 714)	http://mgc.nci.nih.gov/		
	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgapbs-remail.nih.gov			
	Tissue procurement: Jeffrey E. Green, M.D.			
	CDNA Library Preparation: Life Technologies, Inc.			
	CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: LLAM10994	row: c	column: 08	
	High quality sequence stop: 712.			

FEATURES source

```

i. .:714
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:498791"
/lab_host="PH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_C024"
/notes="Organ: colon; Vector: pCMV-SPORT
Site2: Salt; Cloned unidirectionally.
Average insert size: 1.6 kb. Constructed
Technologies, Note: this is a NCI_CGAP
a 178 c 167 g 180 t

```

```

BASE COUNT      189 a      178 c      167 g      180 t
ORIGIN

Query Match      76.88;      Score 334;      DB 12;      Length 714;
Best Local Similarity 86.0%;      Pred. NO. 9.8e-94;
Matches 370;      Conservative 0;      Mismatches 60;      Indels 0;      Gaps 0;

```

1	QY	ATGGGGCCCGCTGCTCAGATTCTTGGGTTCWTGTGCTCTGTTTCCAGGTACCCAGATGT	60
16	Db	ATCAGGGCTCTGCTCAGTTCTTGGCATCTGTTGCTCTGTTTCCAGGTGCCAGATGT	75
61	QY	GACATCCAGATGACCCAGTCTCCATCTCTCTATCTGCCTCTCTGGGCAAAAGAGTCAGT	120
76	Db	GACCTCCAGATGATTCACTTCCATCTCCATGTTGCCTCTCTGGGAGACAGTGTCAGT	135
121	QY	CTCAGTTGTCGGGCAAGTCAGGACATGGGTATTAACTTACATTTGGCTTTCAGCAGGAACCA	180
136	Db	CTCTCTTGTGCGGACTAGTCAGGGCATTAGAGGTAATTTAGACTGGTATCAGCAGAAACCA	195
181	QY	GATGGGAACATTAAACGCCCTCATCTACGCCACATCCAGTTTAGTGTCTGGTCTCCCCAAA	240
196	Db	GGTGGAACTATTAAAGTCCTGATCTACTCCACATCCCAATTTAAATTCGGTGCCTCATCA	255
241	QY	AGGTTTCAGTGGCAGTAGTGTCTGGGTACAGATTATTTCTTCACCATCAGCAGCGCTTGAGTCT	300
256	Db	AGGTTTCAGTGGCAGTGGGTCTGGGTACAGATTATTTCTTCACCATCAGCAGCGCTACAGTCT	315
301	QY	GAAGATTTTGTAGCCATTACTGCTCTACAATATGCTAGTTCTCCGTACACAGTTCGGAGGG	360
316	Db	GAAGATTTTGCAGACTATTACTGCTCTACAAGTGTATGCGTATCCGTGGACGCTTCGGTGGC	375
361	QY	GGGACCAAGCTGGAAATATAAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA	420
376	Db	GGCACCAGCTGGNAATCAAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA	435
421	QY	TCCAGTAAGC	430
436	Db	TCCAGTGAGC	445

RESULT 4  
BY083003  
LOCUS  
DEFINITION

Accession	Length	Source	EST
BY083003	374 bp	mRNA	linear
BY083003	full-length	enriched, pooled tissues, adult spleen,	EST 07-DEC-2002

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

## REFERENCE AUTHORS

etc. Mus musculus cDNA clone K630040M06 5', mRNA sequence.  
 BY083003  
 BY083003.1 GI:26193211  
 EST.  
 Mus musculus (house mouse)

**Mus musculus** (house mouse)

etc. Mus musculus cDNA clone K630040M06 5', mRNA sequence.  
 BY083003  
 BY083003.1 GI:26193211  
 EST.  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 374)

1 (bases 1 to 374)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J.-Y., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojbori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, J. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sanderlin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sulkana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, I. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zvolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kawada, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

I2400851  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

fax: 81-43-503-9240  
Email: genome-res@gsc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane,  
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.,  
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,  
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,  
M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct  
Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

10 (11), 1757-1771 (2000).  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

FEATURES source Location/Qualifiers

1. 374  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="K63040M06"  
/clone\_lib="RIKEN full-length enriched, pooled tissues,  
adult spleen, etc."  
/note="(dev\_stage=adult,tissue\_type=spleen,sex=male),  
(dev\_stage=adult,tissue\_type=kidney,sex=male),  
(dev\_stage=adult,tissue\_type=testis,sex=male),  
(dev\_stage=adult,tissue\_type=thymus,sex=male),  
(dev\_stage=adult,tissue\_type=heart,sex=male),  
(dev\_stage=adult,tissue\_type=colon,sex=male),  
(dev\_stage=adult,tissue\_type=stomach,sex=male),  
(dev\_stage=adult,tissue\_type=liver,sex=male),  
(dev\_stage=13 days embryo,tissue\_type=whole body,sex=mix),  
(dev\_stage=14 days embryo,tissue\_type=whole body,sex=mix),  
(dev\_stage=16 days embryo,tissue\_type=whole body,sex=mix),  
(dev\_stage=17 days embryo,tissue\_type=whole body,sex=mix),  
(dev\_stage=15 days pregnant,adult,tissue\_type=amnion,  
sex=female),(dev\_stage=10 days neonate,tissue\_type=brain,  
sex=mix),(dev\_stage=10 days neonate,tissue\_type=thymus,  
sex=mix),(dev\_stage=10 days neonate,tissue\_type=heart,  
sex=mix)"  
88 a 89 c 85 g 111 t 1 others

BASE COUNT 88 a 89 c 85 g 111 t 1 others  
ORIGIN

Query Match 73.6%; Score 320.2; DB 13; Length 374;  
Best Local Similarity 95.9%; Pred. No. 1.7e-89;  
Matches 326; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
QY 1 ATGGGGCCCTGCTCAGATTCTTGGGCTTCTTGTGCTCTTGTTCCTGCTTCCAGGTACCAGATGT 60  
DB 33 ATGAGGCTCTCTGCACAGATTTTGGGCTTCTTGTGCTCTTGTTCCTGCTTCCAGGTACCAGATGT 92  
QY 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGCGACAAAGAGTCAGT 120  
DB 93 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGCGACAAAGAGTCAGT 152  
QY 121 CTCATTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180  
DB 153 CTCATTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 212  
QY 181 GATGGAACATTTAAACCGCTGATCTACGCCACATPCCAGTTAGTTCTGTTGCCCAAA 240  
DB 213 GATGGAACATTTAAACCGCTGATCTACGCCACATPCCAGTTAGTTCTGTTGCCCAAA 272  
QY 241 AGGTTCAAGTGGCAGTGGCTGAGATTTATCTCTCACCATCAGCAGCCTTGAGTCT 300  
DB 273 AGGTTCAAGTGGCAGTGGCTGAGATTTATCTCTCACCATCAGCAGCCTTGAGTCT 332  
QY 301 GAAGATTTGTAGCTTACTGTCTACAATATGCTAGTTCT 342  
DB 333 GAACATTTGTAGCTTACTGTCTACAATATGCTAGTTCT 374

RESULT 5  
BI152061 602916512f1 NCI\_CGAP\_Lu29 Mus musculus cdna clone IMAGE:5066931 5',  
LOCUS mRNA sequence.  
DEFINITION

BI152061 GI:14612062  
VERSION EST.  
KEYWORDS Mus musculus (house mouse)  
SOURCE  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 827)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11180 row: h column: 04  
High quality sequence stop: 827.

FEATURES

source Location/Qualifiers  
1. 827  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CZECH II"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5066931"  
/tissue\_type="spontaneous tumor, metastatic to mammary.  
Stem cell origin."  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Lu29"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"  
BASE COUNT 235 a 219 c 180 g 193 t  
ORIGIN

Query Match 66.2%; Score 287.8; DB 12; Length 827;  
Best Local Similarity 80.7%; Pred. No. 3.5e-79;  
Matches 348; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 1 ATGGGGCCCTGCTCAGATTCTTGGGCTTCTTGTGCTCTTGTTCCTGCTTCCAGGTACCAGATG 59  
DB 13 ATGATGCTCTGTTCAAGTTCCTTGGGCTCTCTGCTGCTGTTTCAAGGTACCAGATG 72  
QY 60 TGACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGCGACAAAGAGTCAG 119  
DB 73 TGATATCCAGATGACACAGACTACATCTCTGCTCTCTGCGACAAAGAGTCAG 132  
QY 120 TCTCATTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 179  
DB 133 CATCAGTTGAGGCAAGTCAGGACATTTAGATTTAACTGGTATCAGCAAAACC 192  
QY 180 AGATGGAACATTTAAACCGCTGATCTACGCCACATCAGTTAGTTCTGTTGCCCAAA 239  
DB 193 AGATGGAACATTTAAACCGCTGATCTACGCCACATCAGTTAGTTCTGTTGCCCAAA 252  
QY 240 AAGTTCAAGTGGCAGTGGCTGAGATTTATCTCTCACCATCAGCAGCCTTGAGTCT 299  
DB 253 AAGTTCAAGTGGCAGTGGCTGAGATTTATCTCTCACCATCAGCAGCCTTGAGTCT 312  
QY 300 TGAAGATTTGTAGCTTACTGTCTACAATATGCTAGTTCTCCGTACAGTTCGGGAGG 359  
DB 313 AGAAGATTTGCCACITACTTTTGCCAAACAGGATAGTAACTCCGTATAGTTCGATC 372  
QY 360 GGGGACCAAGCTGGAATAAAACGGGCTGATGTCGACCACTGTATCCATCTTCCCAACC 419  
DB 373 GGGGACCAAGCTGGAATAAAACGGGCTGATGTCGACCACTGTATCCATCTTCCCAACC 432  
QY 420 ATCCAGTAAAC 430  
DB 433 ATCCAGTAAAC 443

RESULT 6

BG518527  
LOCUS  
DEFINITION  
BG518527  
ACCESSION  
VERSION

BG518527 874 bp mRNA linear EST 02-APR-2001  
6025782261f1 NCI\_CGAP\_Lu29 Mus musculus cdna clone IMAGE:3491798 5',  
mRNA sequence.  
BG518527  
BG518527.1 GI:13513491



```

QY 361 GGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGATATCCATCTTCCACCA 420
Db 372 GGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGATATCCATCTTCCACCA 431
QY 421 TCCAGTAAGC 430
Db 432 TCCAGTGAGC 441

RESULT 8
BF582283 602101109F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4224267 5',
LOCUS mRNA sequence.
DEFINITION EST 12-DEC-2000
ACCESSION BF582283
VERSION BF582283.1 GI:11655995
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 626)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM9814 row: a column: 04
High quality sequence stop: 624.
FEATURES
Location/Qualifiers
1..626
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4224267"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 178 a 153 c 137 g 158 t
ORIGIN

Query Match 65.6%; Score 285.2; DB 10; Length 626;
Best Local Similarity 80.5%; Pred. No. 2.1e-78;
Matches 346; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

QY 1 ATGGGGCCCCCTGCATCAGATCTTGGGTTCTTGTGCTTCTTCCAGGTACAGATGT 60
Db 1 ATGATGTCCTCTGCTCAGTTCCTTGGTCTCTGCTGCTGCTTCTTCAAGGTACAGATGT 59
QY 61 GACATCCAGATGACCCAGTCTCCATCTCTATCTGCTCTCTGCGGACAAAGAGTCAGT 120
Db 60 GATATCCAGATGACACAGACTACATCCCTCTGCTGCTCTCTGCGGAGACAGATCAC 119
QY 121 CTCACTTGTCCGGCAAGTCAGGACATGGTATTAACCTTACATTTGCTTTCAGCAGAACCA 180
Db 120 ATCAATTCCAGTCAAGTCAGGATTAGCAATTATTAATTGTTATGATCAGCAGAACCA 179
QY 181 GATGGAACATTAACGCTGATCTACGCCACATCCAGTTAGTTTGTGTTGCCCAAA 240
Db 180 GATGGAACATTAACGCTGATCTACGCCACATCCAGTTAGTTTGTGTTGCCCAAA 239
QY 241 AGGTTCACTGCGCAGTAGTCTGGGTCAGATTATCTCTCACCATCAGCAGCCTTGAGTCT 300

```

```

Db 240 AGGTTCACTGCGCAGTAGTCTGGGACAGATTATCTCTCACCATCAGCAGCCTTGAGTCT 299
QY 301 GAAGATTTTGTAGCCTATTACTGCTACAATATGCTAGTCTCCGTACACGTTCCGAGGG 360
Db 300 GAAGATATTGCCACTTACTATTGTCAGCAGTATTTAAGCTTCCATTCAGCGTCCGCTCG 359
QY 361 GGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGATATCCATCTTCCACCA 420
Db 360 GGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGATATCCATCTTCCACCA 419
QY 421 TCCAGTAAGC 430
Db 420 TCCAGTGAGC 429

RESULT 9
BF580940 602100636F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4224032 5',
LOCUS mRNA sequence.
DEFINITION EST 12-DEC-2000
ACCESSION BF580940
VERSION BF580940.1 GI:11654652
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 725)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM9813 row: g column: 09
High quality sequence stop: 694.
FEATURES
Location/Qualifiers
1..725
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4224032"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 207 a 183 c 159 g 176 t
ORIGIN

Query Match 65.6%; Score 285.2; DB 10; Length 725;
Best Local Similarity 80.5%; Pred. No. 2.2e-78;
Matches 346; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

QY 1 ATGGGGCCCCCTGCATCAGATCTTGGGTTCTTGTGCTTCTTCCAGGTACAGATGT 60
Db 20 ATGATGTCCTCTGCTCAGTTCCTTGGTCTCTGCTGCTGCTTCTTCAAGGTACAGATGT 78
QY 61 GACATCCAGATGACCCAGTCTCCATCTCTATCTGCTCTCTGCGGACAAAGAGTCAGT 120
Db 79 GATATCCAGATGACACAGACTACATCCCTCTGCTGCTCTCTGCGGAGACAGATCAC 138
QY 121 CTCACTTGTCCGGCAAGTCAGGACATGGTATTAACCTTACATTTGCTTTCAGCAGAACCA 180

```

```

Db      139 ATCAGTTGCGAGTGAAGTCAGGCAATTAGTAATATTATTAACATGGTATATCAGCAAAACCA 198
Qy      181 GATGGAACTATTAAACCCCTGATCTAGCCACATCCAGTTTAGTCTTCTGCTGTCGCCAAA 240
Db      199 GATGGAACTTTAACTCTGATCTATTACACATCAAGTTTACACTCAGGAGTCCCATCA 258
Qy      241 AGTTTCAGTGGCAGTAGTGGTCTGGGTGAGATTTATCTCTACCATCAGCAGCCCTTGAGTCT 300
Db      259 AGTTTCAGTGGCAGTAGTGGTCTGGGACACATTTATCTCTACCATCAGCAACCTGGAACCT 318
Qy      301 GAAGATTTTTCAGCCATTACTCTACAAATATATGCTAGTTCTCGGTACACGTTTCGGAGGG 360
Db      319 GAAGATATTGCCCTTACTATTCTAGCAGTATAGTTCATTCATTCAGTTTCGGGTCTG 378
Qy      361 GGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGTATCTATCTTCCCAACCA 420
Db      379 GGCACAAAGTTGGAATAAAACGGGCTGATGCTGCACCAACTGTATCTTCCCAACCA 438
Qy      421 TCCAGTAAGC 430
Db      439 TCCAGTGAGC 448

RESULT 10
BQ918407
LOCUS      1459 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT_8885602 NCI_CGAP_Co24 Mus musculus cDNA clone
IMAGE: 6397382 5', mRNA sequence.
ACCESSION BQ918407.1 GI:22333105
VERSION BQ918407.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1459)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13895 row: c column: 15
High quality sequence start: 165
High quality sequence stop: 634.
FEATURES
source
1. .1459
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="FVB/N"
/clone="IMAGE:6397382"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 393 a 441 c 306 g 315 t 4 others
ORIGIN
Query Match 65.0%; Score 282.8; DB 13; Length 1459;
Best Local Similarity 78.6%; Pred. No. 1.6e-77;
Matches 338; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
Qy      1 ATGGGGCCCTGCTCAGATCTCTGGTCTTCTGCTCTGTTTCCAGGTACCAATGT 60

```

```

Db      190 ATGATGTCTCTGCTCAGTTCTCTGCTCTCTGTTCTCTGTTTCTAGGTTCAGATGT 249
Qy      61 GACATCCAGATGACCCAGTCTCCATCTCTCTTATCTGCTCTCTCTGGGACAAAGAGTCAGT 120
Db      250 GAAATCCAGATGACACAGACTACATCTCTCTGCTCTCTCTGCGGACAGAGTCAGTCC 309
Qy      121 CTCACATTGTCTGGGCAAGTCAGGACATTTGGTATTAACTTTACATTTGGTCTTCCAGGAACCA 180
Db      310 ATCATTGTGAGGTCAAGTCAGGACATTTAGTAAATTTATTTAACTGGTATCAGCAAAACCA 369
Qy      181 GATGGAACTATTAAACCCCTGATCTAGCCACATCCAGTTTAGTCTTCTGTTGTTCCCAAA 240
Db      370 GATGGAACTATTAAACCTCTGATCTACTACACATCAAAATATACACTCAGAGTCCCATCA 429
Qy      241 AGTTTCAGTGGCAGTAGTGGTCTGGGTGAGATTTATCTCTCACCATTACCACTGAGGACAA 489
Db      430 AGTTTCAGTGGCAGTAGTGGTCTGGACAGATTTATCTCTCACCATTACCACTGAGGACAA 489
Qy      301 GAAGATTTTTCAGCTATTACTGTCTACAATATGCTAGTTCTCCGTACAGCTTCCGAGGG 360
Db      490 GAGGATATTGCCACTTACTTTTGCCAAACAGGTAATACACTTCCGTGGACGTTCCGTTGGA 549
Qy      361 GGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGTATCTTCCATCTTCCCAACCA 420
Db      550 GGCACCAAACTAGAAAGTCAATCGGCTGATGCTGCACCAACTGTATCTTCCATCTTCCCAACCA 609
Qy      421 TCCAGTAAGC 430
Db      610 TCCAGTGAGC 619

RESULT 11
BQ144806
LOCUS      762 bp mRNA linear EST 24-OCT-2000
DEFINITION BQ1791486F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4022181 5',
mRNA sequence.
ACCESSION BQ144806
VERSION BQ144806.1 GI:10983846
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 762)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Arrayed by: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9278 row: 1 column: 22
High quality sequence stop: 677.
FEATURES
source
1. .762
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4022181"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; transgenic model WNT-1, expression driven by
MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dt. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

```



```

BASE COUNT      215 a      182 c      183 g      182 t
ORIGIN

Query Match      64.1%; Score 278.8; DB 10; Length 762;
Best Local Similarity 78.4%; Pred. No. 2.3e-76;
Matches 334; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 5 GGGCCCTGCTCAGATCTTGGGCTCTTGGTCTTGTTCCTGTTTCCAGGTACCAAGTGTGACA 64
Db 1 GGACCCCTGCTCAGTGGGCTTGGATCTTGTGCTCTGCTGCTTCCAGGTATCACTGTGACA 60
Qy 65 TCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGTCACTCTCA 124
Db 61 TCAAGATGACCCAGTCTCCATCTTCCATGTATGATCTCTAGGGAGAGAGTCACTATCA 120
Qy 125 CTTCTGGGCAAGTCAGACATTTGGTATTAATTAATTTACATTTGCTTCAGCAGGAACCAAGT 184
Db 121 CTTCAAGGCGAGTCAGACATTAATAGCTATTTACATTTGCTTACCAGCAGAAACCAAGAA 180
Qy 185 GAATCTAATAAGGCTGATCTACGCCACATCCAGTTTGGTTCCTGCTGCTCCCAAAAGGT 244
Db 181 AATCTCTAAGACCTGATCTATGCTGCAACACAGCTTGGTAGATGGGTCCTCCATCAAGGT 240
Qy 245 TCAGTGGCAGTGTCTGGGTCAGATTTCTCTCACCATCAGCAGCTTGGAGTCTGAAG 304
Db 241 TCAGTGGCAGTGTCTGGGTCAGATTTCTCTCACCATCAGCAGCTTGGAGTCTGAAG 300
Qy 305 ATTTTGTAGCCTTACTGCTACAAATATGCTAGTTCTCTCCGTACACAGCTTGGAGTCTGAAG 364
Db 301 ATATAGGAATTTATTTATTTGCTACAGATGATGATGTTCCGTATACGTTCCGATCGGGGA 360
Qy 365 CCAAGCTGGAATAAAGCGGCTGATGCTGCACCACTGTATCATCTTCCCAACCACTCA 424
Db 361 CCAAGCTGGAATAAAGCGGCTGATGCTGCACCACTGTATCATCTTCCCAACCACTCA 420
Qy 425 GTAGGC 430
Db 421 GTAGGC 426

RESULT 12
LOCUS      B0943305
DEFINITION AGENCOURT_8880809 NCI_CGAP_Co24 Mus musculus cDNA clone EST 21-AUG-2002
IMAGE:6397142 5', mRNA sequence.
ACCESSION  B0943305
VERSION    B0943305.1 GI:22358783
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE     NIH-MGC http://mgi.nci.nih.gov/
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: The Cepko Laboratory
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone Distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L1AM13894 row: 1 column: 15
          High quality sequence stop: 670.
          Location/Qualifiers
            1..901
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="FVB/N"
              /db_xref="taxon:10090"

/clone="IMAGE:6397142"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1; NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      231 a      243 c      199 g      227 t
ORIGIN

Query Match      63.9%; Score 278; DB 13; Length 901;
Best Local Similarity 77.9%; Pred. No. 4.4e-76;
Matches 335; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 1 ATGGGGCCCTGCTCAGATCTTGGGCTCTTGTGCTTGTTCCTGTTTCCAGGTACCAAGTGT 60
Db 14 ATGAGGACCCCTGCTCAGATCTTGGGCTCTTGTGCTTGTTCCTGTTTCCAGGTGTCAAAATGT 73
Qy 61 GACATCCAGATGACCCAGTCTCCATCTCTCTATCTGCTCTCTGGGACAAAGAGTCACT 120
Db 74 GACATCAAGATGACCCAGTCTCCATCTTCCATGTATGATCTCTAGGAGAGGCGGTCACT 133
Qy 121 CTCATCTTGTGGGCAAGTCAGGACATTTGGTATTAATTAATTTACATTTGCTTCCAGCAGGAACCA 180
Db 134 ATCACTTGCAGGCGAGTCAGGACATTTATTCCTTATTAATTTGTTCCAGCAGGAACCA 193
Qy 181 GATGGAATATTAACGCCCTGATCTACGCCACATCCAGTTTGGTTCCTGCTGCTCCCAAA 240
Db 194 GGAAGCCTCTCTAAGACCTGATCTATCATCAACACAGATTTGATGGGTCCTCCCTCA 253
Qy 241 AGTTTCAGTGGCAGTGTCTGGGTCAGATTTATTCCTCACCATCAGCAGCTTGGAGTCT 300
Db 254 AGTTTCAGTGGCAGTGTCTGGGTCAGATTTATTCCTCACCATCAGCAGCTTGGAGTCT 313
Qy 301 GAAGATTTGTAGCCTTACTGCTACAAATATGCTAGTTCTCCGTACACCTTCCGAGGG 360
Db 314 GAAGATTTGTAGCCTTACTGCTACAAATATGCTAGTTCTCCGTACACCTTCCGAGGG 373
Qy 361 GGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCACTGTATCATCTTCCCAACCA 420
Db 374 GGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCACTGTATCATCTTCCCAACCA 433
Qy 421 TCCAGTAAAGC 430
Db 434 TCCAGTAAAGC 443

RESULT 13
LOCUS      BF581992
DEFINITION 602099448F1 NCI_CGAP_Co24 Mus musculus cDNA clone EST 12-DEC-2000
IMAGE:6397142 5', mRNA sequence.
ACCESSION  BF581992
VERSION    BF581992.1 GI:11655613
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE     NIH-MGC http://mgi.nci.nih.gov/
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone Distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L1AM9801 row: 0 column: 11

```



High quality sequence stop: 715.

## FEATURES

source  
1. .772  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4219618"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCL\_CGAP\_Co24"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dr.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 216 a 204 c 175 g 176 t 1 others  
ORIGIN

Query Match 63.5%; Score 276.4; DB 10; Length 772;  
Best Local Similarity 80.7%; Pred. No. 1.3e-75;  
Matches 322; Conservative 0; Mismatches 77; Indels 0; Gaps 0;  
QY 32 TGTGCTCTGTTTCCAGGTACCAGATGTGACATCCAGATGACCCAGTCTCCATCTCCT 91  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 30 TGTGCTCTGTTTCAAGTTCAGATGTGATATCCAGATGACACAGACTACATCTCC 89  
QY 92 TATCTGCCCTCTGGGACAAAGAGTCAGTCTCACTTGTGGGCAAGTCAGGACATTGGTA 151  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 90 TGTCTGCCCTCTGGGAGACAGATCACCATCAGTTGCAGTGGAACTCAGGGCATGGCA 149  
QY 152 TTAACATTACATGGCTTCAGCAGGACACAGATGGAACATTAACGCCCTGATCAGCCA 211  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 150 ATTATTAACTGGTATCAGCAAAACAGATGGAACCTGTTAAACTCTCTGATCTATACA 209  
QY 212 CATCCAGTTAGTGTCTGCTGTCCTCCCAAGGTTCACTGGCAGTAGTCTGGGTCAAGTT 271  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 210 CATCAAGTTACACTCAGGAGTCCCATCAAGTTCACTGGCATGGTCTGGGACAGATT 269  
QY 272 ATTCTCTACCATCAGCAGCTTGAGTCTGAAGATTTGTAGCCTATTACTGTCTCAAT 331  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 270 ATTCTCTACCATCAGCAACCTGGAACCTGAAGATATNGCCACTTACTATTGTTCAGCAGT 329  
QY 332 ATCTAGTTCTCGGTACAGTTCGGAGGGGGACCAAGCTGGAATTAACAGGGCTGATG 391  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 330 ATCGTTACCTTCGTGGAGCTTCGGTGGAGGGACCAAGTTGGAATCAACAGGGCTGATG 389  
QY 392 CTGCACCAACTGTATCCATCTTCCCACTCCAGTAAGC 430  
Db 390 CTGCACCAACTGTATCCATCTTCCCACTCCAGTGAGC 428

## RESULT 14

BI455041 830 bp mRNA linear EST 21-AUG-2001  
LOCUS 603173343F1 NCL\_CGAP\_Mam5 Mus musculus cdna clone IMAGE:5252723 5',  
DEFINITION mRNA sequence.

ACCESSION BI455041 GI:15245697

VERSION BI455041.1

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 830)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11638 row: e column: 12

High quality sequence stop: 820.

## FEATURES

source

1. .830  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5252723"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/clone\_lib="NCL\_CGAP\_Mam5"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dr.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"  
BASE COUNT 227 a 230 c 172 g 201 t  
ORIGIN

Query Match 63.5%; Score 276.4; DB 12; Length 830;

Best Local Similarity 77.7%; Pred. No. 1.4e-75;

Matches 334; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 1 ATGGGGCCCTCCTCAGATTCCTGGTCTCTGCTCTGCTCTGTTTCCAGGTACCAGATGT 60  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 6 ATGATGCTCTCTCAGTTCCTGGTCTCTGCTCTGCTCTGCTCTCAAGGTTCAGATGT 65  
QY 61 GACATCCAGATGACCCAGTCTCCATCTCTCTATCTGCTCTCTGGGACAAAGAGTCAGT 120  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 66 GATATCCAGATGACCCAGACTACATCTCTCTGCTCTCTGGGGACAGAGTCACC 125  
QY 121 CTCACITGTCGGCAGTCAGGACATTTGATTAACATTACATTGCTTCCAGCAGGAACCA 180  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 126 ATCAATTTGAGGCAAGTCAGGACATTAACCAATTTTAAATTTGGTATCAGAGAAACCA 185  
QY 181 GATGGAACTATTAAAGCCCTGATCTAGCCACATCCAGTTTAGTCTCTGCTGCCCAAA 240  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 186 AATGGAACTTTTAAATCTCTGATCTCTCTACACTTCCATATTACACTCAGAGTCCCATCA 245  
QY 241 AGTTTCAGTGGCAGTAGGTCTGGTCTCAGATTTATTTCTTCCATCAGCAGCCCTTGAGTCT 300  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 246 AGTTTCAGTGGCAGTGGTCTGGAACAGATTTATTTCTCACCATTAGCAACCTTGAACGA 305  
QY 301 GAAGATTTGTAGCCTATTACTGCTTACAATATGCTAGTCTCGGTACACGTTCCGAGGG 360  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 306 GAAGATATTGCCACTTACTTTTCCCAACAGGGTAATTCGCTCTCTCAGCTTCGGTGTCT 365  
QY 361 GGGACCAAGCTGGAATAAAACGGGTGATGCTGCACCACTGATTCATCTTCCATCTTCCCAACCA 420  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 366 GGGACCAAACTGGAATGAACGGGTGATGCTGCACCACTGATGCTGCACCACTGATTCATCTTCCCAACCA 425  
QY 421 TCCAGTAAGC 430  
Db 426 TCCAGTGAGC 435

## RESULT 15

BE284224

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 695)

BE284224 695 bp mRNA linear EST 13-JUL-2000  
LOCUS 601099161F1 NCL\_CGAP\_Lu29 Mus musculus cdna clone IMAGE:3491798 5',  
DEFINITION mRNA sequence.

ACCESSION BE284224

VERSION BE284224.1

KEYWORDS GI:9161016

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 695)

Search completed: August 30, 2003; 19:44:46  
Job time : 1519.79 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 22:04:11 ; Search time 24.8154 Seconds  
(without alignments)  
927.461 Million cell updates/sec

Title: US-08-836-455-2

Perfect score: 145

Sequence: 1 MGAPQILGFLLLFPGRRC.....IKRDAAPVTSTPPSKLG 145

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145	100.0	145	18	AAW27119 Murine monoclonal
2	145	100.0	145	24	AAQ16292 Mouse 11D10 antibo
3	90	62.1	145	20	AAW87593 Antibody 11D10 lig
4	33	22.8	129	18	AAW22537 Murine anti-human
5	33	22.8	129	23	AAW20200 Murine 44H104 mab
6	33	22.8	129	23	AAE20204 Murine 44H104 mab
7	30	20.7	146	16	AAW74966 Anti-idiotypic anti
8	30	20.7	206	20	AAW39452 Antibody ABX-CBL 1
9	29	20.0	92	20	AAW95480 Mouse derived RT3

10	29	20.0	92	20	AAW95476	Mouse derived RT3
11	29	20.0	95	19	AAW70954	Light chain of mon
12	29	20.0	107	21	AAW90824	520C9 hybridoma VL
13	29	20.0	108	13	AAW21310	Light chain of Mif
14	29	20.0	108	21	AAW44587	Mouse anti-IL-18 a
15	29	20.0	130	24	AAE34366	Escherichia coli 1
16	29	20.0	135	21	AAW44599	Mouse light chain
17	29	20.0	146	4	AAW30251	Sequence of the le
18	29	20.0	149	17	AAW03199	Anti-idiotypic mono
19	29	20.0	149	20	AAW49209	MAB 1A7 light chai
20	29	20.0	149	20	AAW28468	Light chain variab
21	29	20.0	149	20	AAW21545	Monoclonal antibod
22	29	20.0	237	21	AAW44595	EscFv#125-2H recom
23	29	20.0	243	17	AAW02280	520C9 anti-c-erbB-
24	29	20.0	243	19	AAW53170	520C9 anti-c-erbB-
25	29	20.0	243	20	AAW80424	520C9 sFv sequence
26	29	20.0	243	21	AAW44596	EscFv#125-2H HT te
27	29	20.0	246	14	AAW39569	Sequence of 520C9
28	29	20.0	252	23	AAW72863	P4-14 single chain
29	29	20.0	257	23	AAW72869	P5-11 single chain
30	29	20.0	499	23	AAW72872	3B10xP4-14 bispeci
31	29	20.0	534	14	AAW39571	Sequence of G-FIT
32	28	19.3	107	21	AAW90825	650E2 hybridoma VL
33	28	19.3	112	23	AAW50879	Anti-bovine prion
34	28	19.3	148	22	AAW35096	Filamentous phage
35	28	19.3	213	10	AAW93035	Chimeric antibody
36	28	19.3	214	18	AAW27089	Mouse monoclonal a
37	28	19.3	216	17	AAW15935	Antibody 3G2 light
38	28	19.3	218	16	AAW75457	Mouse antibody F83
39	28	19.3	218	16	AAW75459	Mouse antibody F4-
40	28	19.3	218	24	ABP96753	TSH receptor antib
41	28	19.3	218	24	ABP96754	TSH receptor antib
42	28	19.3	219	21	AAW95258	WOW-1 Fab light ch
43	28	19.3	219	21	AAW68994	Amino acid sequenc
44	28	19.3	234	23	ABG31731	Murine m166 antibo
45	28	19.3	240	12	AAW15443	Single chain Fv fr

#### ALIGNMENTS

RESULT 1  
AAW27119  
ID AAW27119 standard; Protein; 145 AA.

AC AAW27119;

DT 25-MAR-2003 (updated)

DT 04-JAN-1998 (first entry)

XX Murine monoclonal anti-idiotypic antibody 11D10 VL region.

DE Monoclonal antibody 11D10; anti-idiotypic antibody; mucin;  
XX human milk fat globule; HWFG; tumour; breast cancer; vaccine.

XX Mus musculus.

XX Key Location/Qualifiers

PH Peptide 1..20

FT /label= Sig\_peptide

FT Protein 21..145

FT /label= Mat\_protein

FT Region 21..43

FT /label= FR1

FT /note= "framework region 1"

FT Region 44..54

FT /label= CDR1

FT /note= "complementarity determining region 1"

FT Region 55..69

FT /label= FR2

FT /note= "framework region 2"

FT Region 70..76

FT /label= CDR2

```

FT FT /note= "complementarity determining region 2"
FT FT 77..108
FT FT /label= FR3
FT FT /note= "framework region 3"
FT FT 109..117
FT FT /label= CDR3
FT FT 118..127
FT FT /note= "complementarity determining region 3"
FT FT /label= FR1
FT FT /note= "framework region 4"
XX XX
FN W09722699-A2.
XX XX
XX PD 26-JUN-1997.
XX XX
XX PF 19-DEC-1996; 96WO-US20757.
XX XX
XX PR 20-DEC-1995; 95US-0575762.
XX PR 26-JAN-1996; 96US-0591965.
XX PR 13-DEC-1996; 96US-0766350.
XX XX
XX PA (KENT ) UNIV KENTUCKY.
XX XX
XX PI Chatterjee M, Chatterjee SK, Foon KA;
XX XX
XX DR WPI; 1997-341690/31.
XX DR N-PSDB; AAT85149.
XX XX
XX PT Monoclonal anti-idiotype antibody 11D10 - elicits immune response
XX PT against human milk fat globule disease associated tumours,
XX PT especially breast cancer
XX PS Claim 9; Page 94; 130pp; English.
XX XX
XX CC This polypeptide sequence comprises the light chain variable region
XX CC (VL) of monoclonal anti-idiotype antibody 11D10 produced by
XX CC hybridoma cell line ATCC 12020. 11D10 was obtained by immunising
XX CC naive mice with MC-10 anti-HMFG antibody to obtain an anti-idiotype
XX CC response. It elicits an immune response against a specific epitope
XX CC of a high mol.wt. mucin of human milk fat globule (HMFG). It
XX CC induces an immunological response to HMFG in mice, rabbits, monkeys
XX CC and patients with advanced HMFG-associated tumours. Pharmaceutical
XX CC compositions and vaccines comprising 11D10, 11D10 polypeptides
XX CC and/or 11D10 polynucleotides (see also AAT85149-50) are claimed.
XX CC Also claimed are diagnostic kits and methods of using 11D10, 11D10
XX CC polypeptides and/or 11D10 polynucleotides, including methods of
XX CC treating HMFG-associated tumours. 11D10 is also used in a claimed
XX CC method of palliating HMFG-associated disease and in claimed kits to
XX CC detect or quantify anti-HMFG antibody.
XX CC (Updated on 25-MAR-2003 to correct PR field.)
XX XX
SQ Sequence 145 AA;
Query Match 100.0%; Score 145; DB 18; Length 145;
Best Local Similarity 100.0%; Pred. No. 5.7e-125;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGAPAIQLGFLLLFPGTRCDIQMTQSPSSLSASLGQVSLTCSRASQDIGINLHWLQOEP 60
DB 1 MGAPAIQLGFLLLFPGTRCDIQMTQSPSSLSASLGQVSLTCSRASQDIGINLHWLQOEP 60
QY 61 DGTIKRLIYATSSLGSGVPRKFRSGSRGSDYSLTIFSSLEDFVAYYCLQYASSPYTFGG 120
DB 61 DGTIKRLIYATSSLGSGVPRKFRSGSRGSDYSLTIFSSLEDFVAYYCLQYASSPYTFGG 120
QY 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145
DB 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145
RESULT 2
AAO16292
ID AAO16292 standard; Protein; 145 AA.
XX XX
XX AC AAO16292;
XX XX
XX DT 20-MAR-2003 (first entry)
XX XX
XX DE Mouse 11D10 antibody light chain variable region.
XX XX
XX KW Mouse; murine; vaccine; tumour; human milk fat globules; HMFG;
XX KW carcinoembryonic antigen; CEA; 11D10; 3H1; HMFG-associated tumour;
XX KW CEA-associated tumour; anti-idiotype antibody.
XX OS Mus musculus.
XX XX
XX FH Key Location/Qualifiers
XX FT Peptide 1..20
XX FT /label= signal_peptide
XX FT Protein 21..145
XX FT /note= "Mature murine 11D10 antibody light chain variable
XX FT region"
XX FN W0200292012-A2.
XX XX
XX PD 21-NOV-2002.
XX XX
XX PF 17-MAY-2002; 2002WO-US15840.
XX XX
XX PR 17-MAY-2001; 2001US-0861294.
XX XX
XX PA (KENT ) UNIV KENTUCKY RES FOUND.
XX XX
XX PI Chatterjee M, Foon KA;
XX XX
XX DR WPI; 2003-129216/12.
XX DR N-PSDB; AAL51273.
XX XX
XX CC Use of anti-idiotypic antibodies for human milk fat globules (HMFG)- or
XX CC carcinoembryonic antigen (CEA)-associated tumor for delaying the
XX CC development of, or treating a HMFG- or CEA-associated tumor (e.g.
XX CC breast tumor) in humans
XX PS Claim 2; Fig 1; 98pp; English.
XX CC The invention comprises a method for delaying the development of, or
XX CC treating a tumor that is associated with human milk fat globules (HMFG)
XX CC or carcinoembryonic antigen (CEA). The method of the invention involves
XX CC administering an anti-idiotypic antibody for HMFG (e.g. 11D10), and an
XX CC anti-idiotypic antibody for CEA (e.g. 3H1). The method is useful for
XX CC delaying the development, of or treating HMFG/CEA-associated tumours. The
XX CC present amino acid sequence represents the light chain variable region of
XX CC the mouse 11D10 anti-idiotype antibody.
XX XX
SQ Sequence 145 AA;
Query Match 100.0%; Score 145; DB 24; Length 145;
Best Local Similarity 100.0%; Pred. No. 5.7e-125;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGAPAIQLGFLLLFPGTRCDIQMTQSPSSLSASLGQVSLTCSRASQDIGINLHWLQOEP 60
DB 1 MGAPAIQLGFLLLFPGTRCDIQMTQSPSSLSASLGQVSLTCSRASQDIGINLHWLQOEP 60
QY 61 DGTIKRLIYATSSLGSGVPRKFRSGSRGSDYSLTIFSSLEDFVAYYCLQYASSPYTFGG 120
DB 61 DGTIKRLIYATSSLGSGVPRKFRSGSRGSDYSLTIFSSLEDFVAYYCLQYASSPYTFGG 120
QY 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145
DB 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145
RESULT 3
AAW87593
ID AAW87593 standard; Protein; 145 AA.

```

XX AC AAW87593;  
 XX DT 16-MAR-1999 (first entry)  
 XX DE Antibody 11D10 light chain variable region.  
 XX KW Murine; mouse; antibody; light chain; variable region; anti-idiotypic;  
 KW human milk fat globule; tumour; ovary; lung; pancreas; carcinoma; breast.  
 XX OS Mus sp.  
 XX PH Location/Qualifiers  
 FT Peptide 1..20  
 FT Protein /note= "signal peptide"  
 FT Region 21..999  
 FT Region /note= "mature protein"  
 FT Domain 21..43  
 FT Domain /label= framework\_1  
 FT Region 44..54  
 FT Region /label= CDR1  
 FT Region /note= "complementarity determining region 1"  
 FT Region 55..69  
 FT Region /label= framework\_2  
 FT Misc-difference 55  
 FT Domain /note= "encoded by TGG"  
 FT Domain 70..76  
 FT Region /label= CDR2  
 FT Region /note= "complementarity determining region 2"  
 FT Domain 77..108  
 FT Region /label= framework\_3  
 FT Domain 109..117  
 FT Region /label= CDR3  
 FT Region /note= "complementarity determining region 3"  
 FT Region 118..127  
 FT Region /label= framework\_4  
 XX WO9856419-A1.  
 XX PN 17-DEC-1998.  
 XX PD 12-JUN-1998; 98WO-US12250.  
 XX PF 11-JUN-1998; 98US-0096244.  
 XX PR 13-JUN-1997; 97US-0049540.  
 XX PS (KENT ) UNIV KENTUCKY RES FOUND.  
 XX PI Chatterjee M, Poon KA;  
 XX WIPI; 1999-060029/05.  
 XX N-PSDB; AAV83772.  
 XX PT Delaying development of, or treating, HMFG-associated tumours -  
 PT using anti-idiotypic antibody 11D10 raised against antibodies to  
 PT human milk fat globule protein  
 XX PS Disclosure; Fig 1; 54pp; English.  
 XX CC This sequence represents the light chain variable region of the murine  
 CC antibody 11D10. This anti-idiotypic antibody is used to delay the  
 CC development of, or treat, a human milk fat globule (HMFG) associated  
 CC tumour in an individual having low tumour burden. The antibody 11D10  
 CC is used to prevent the recurrence of HMFG-associated tumours e.g.  
 CC ovarian, non-small cell lung and pancreatic carcinoma, especially for  
 CC treating breast tumours.  
 XX SQ Sequence 145 AA;  
 Query Match 62.1%; Score 90; DB 20; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-74;  
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 56 LQEPDGTIKRLIYATSSLGSGVPKRFGSGRSGDYSLTISSESEDFVAYICQYASSP 115  
 DB 56 LQEPDGTIKRLIYATSSLGSGVPKRFGSGRSGDYSLTISSESEDFVAYICQYASSP 115  
 QY 116 YTEGGGTKEIKRADAAPTVSIIPPSSKLG 145  
 DB 116 YTEGGGTKEIKRADAAPTVSIIPPSSKLG 145  
 RESULT 4  
 AAW22537  
 ID AAW22537 standard; Protein; 129 AA.  
 XX AC AAW22537;  
 XX DT 03-NOV-1997 (first entry)  
 XX DE Murine anti-human class II monoclonal antibody 44H104 VL chain.  
 XX KW Antibody; light chain; variable region; hybridoma cell line 44H104;  
 KW immune response; enhance; stimulate; vaccine; immunodiagnosis;  
 KW antigen delivery.  
 XX OS Mus musculus.  
 XX PN WO9640941-A1.  
 XX PD 19-DEC-1996.  
 XX PF 07-JUN-1996; 96WO-CA00400..  
 XX PR 07-JUN-1995; 95US-0483576.  
 XX PA (CONN-) CONNAUGHT LAB LTD.  
 XX PI Anand NN, Barber BH, Caterini JE, Cates GC, Klein MH;  
 XX WIPI; 1997-077271/07.  
 XX N-PSDB; AAT77851.  
 XX PT Recombinant conjugate antibody mol., modified for delivering an  
 PT antigen - elicits enhanced immune response without the use of  
 PT adjuvant to generate antibodies which are useful in vaccines or  
 PT immuno:diagnosis  
 XX PS Example 1; Fig 1A; 64pp; English.  
 XX CC Novel recombinant conjugate antibody molecules comprise a monoclonal  
 CC antibody specific for a surface structure of antigen presenting  
 CC cells (APC), genetically modified to contain at least one antigen  
 CC exclusively at one or more preselected sites. The conjugate is capable  
 CC of delivering the antigen to APC and eliciting an immune response to  
 CC the antigen. The new conjugates are useful as vaccines and are able  
 CC to elicit an enhanced immune response without the use of an adjuvant.  
 CC In a specific example, a conjugate was constructed using the murine  
 CC anti-human class II monoclonal antibody secreted by hybridoma  
 CC 44H104. The peptide CLTB36 was chosen as antigen; it consists of  
 CC a tandemly linked T and B cell epitope derived from HIV MN strain.  
 CC The present sequence represents the light chain variable region  
 CC from 44H104 which was used in the preparation of a conjugate with  
 CC antigen CLTB36.  
 XX SQ Sequence 129 AA;  
 Query Match 22.8%; Score 33; DB 18; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-22;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 FPGTRCDIQMTQSPSSLSASLGQRVSLTCRASQ 47  
 DB 17 FPGTRCDIQMTQSPSSLSASLGQRVSLTCRASQ 49

## RESULT 5

AAE20200  
ID AAE20200 standard; Protein: 129 AA.

XX AC AAE20200;

XX DT 18-JUN-2002 (first entry)

XX DE Murine 44H104 mab variable light chain (VL).

XX KW Murine; mab; light chain; VL; conjugate antibody; antigen delivery;  
XX KW immune system; vaccine; detecting agent; antibacterial.

XX OS Mus sp.

XX PN US2002025315-A1.

XX PD 28-FEB-2002.

XX PF 14-JAN-1998; 98US-0007093.

XX PR 14-JAN-1998; 98US-0007093.

XX PA (ANAN/) ANAND N N.

XX PA (BARB/) BARBER B H.

XX PA (CATE/) CATES G A.

XX PA (CATE/) CATERINI J E.

XX PA (KLEI/) KLEIN M H.

XX PI Anand NN, Barber BH, Cates GA, Caterini JE, Klein MH;

XX DR WPI; 2002-267519/31.

XX DR N-PSDB; AAD32138.

XX PT Novel recombinant conjugate antibody, useful as a vaccine against  
XX PT pathogens having a specific antigen, comprises a monoclonal antibody  
XX PT specific for an antigen presenting cell surface structure -

XX PS Example 1; Fig 1A; 28pp; English.

XX CC The invention relates to a recombinant conjugate antibody, comprising  
XX CC a monoclonal antibody specific for a surface structure of antigen  
XX CC presenting cells genetically modified to contain an antigen moiety  
XX CC for the purpose of delivery of the antigen moiety to antigen-  
XX CC presenting cells of the immune system. The conjugate antibody is  
XX CC formulated as a vaccine to protect a host against a disease caused  
XX CC by a pathogen expressing the antigen. The antibody is useful as a  
XX CC detecting agent. The present sequence is murine 44H104 mab variable  
XX CC light chain (VL).

XX CC Note: This sequence SEQ.ID.NO:2 is stated to be similar to the  
XX CC sequence shown in page 11-12 of the specification (AAE20204).

XX CC However these sequences differ.

XX SQ Sequence 129 AA;

Query Match

Best Local Similarity 22.8%; Score 33; DB 23; Length 129;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 FPGTRCDIQMTQSPSSLSASLGQRVSLTCRASQ 47

DB 17 FPGTRCDIQMTQSPSSLSASLGQRVSLTCRASQ 49

## RESULT 6

AAE20204  
ID AAE20204 standard; Protein: 129 AA.

XX AC AAE20204;

XX DT 18-JUN-2002 (first entry)

XX DE Murine 44H104 mab variable light chain (VL), alternative version.

XX

KW Murine; mab; light chain; VL; conjugate antibody; antigen delivery;  
KW immune system; vaccine; detecting agent; antibacterial.

XX OS Mus sp.

XX FH Key Location/Qualifiers

XX FT Misc-difference 93.

XX FT /note= "Encoded by TAT"

XX PN US2002025315-A1.

XX PD 28-FEB-2002.

XX PF 14-JAN-1998; 98US-0007093.

XX PR 14-JAN-1998; 98US-0007093.

XX PA (ANAN/) ANAND N N.

XX PA (BARB/) BARBER B H.

XX PA (CATE/) CATES G A.

XX PA (CATE/) CATERINI J E.

XX PA (KLEI/) KLEIN M H.

XX PI Anand NN, Barber BH, Cates GA, Caterini JE, Klein MH;

XX DR WPI; 2002-267519/31.

XX DR N-PSDB; AAD32138.

XX PT Novel recombinant conjugate antibody, useful as a vaccine against  
XX PT pathogens having a specific antigen, comprises a monoclonal antibody  
XX PT specific for an antigen presenting cell surface structure -

XX PS Example 1; Page 11-12; 28pp; English.

XX CC The invention relates to a recombinant conjugate antibody, comprising  
XX CC a monoclonal antibody specific for a surface structure of antigen  
XX CC presenting cells genetically modified to contain an antigen moiety  
XX CC for the purpose of delivery of the antigen moiety to antigen-  
XX CC presenting cells of the immune system. The conjugate antibody is  
XX CC formulated as a vaccine to protect a host against a disease caused  
XX CC by a pathogen expressing the antigen. The antibody is useful as a  
XX CC detecting agent. The present sequence is murine 44H104 mab variable  
XX CC light chain (VL), alternative version.

XX CC Note: This sequence SEQ.ID.NO:2 is stated to be similar to the  
XX CC sequence shown in fig 1A (AAE20200). However these sequences differ.

XX SQ Sequence 129 AA;

Query Match

Best Local Similarity 22.8%; Score 33; DB 23; Length 129;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 FPGTRCDIQMTQSPSSLSASLGQRVSLTCRASQ 47

DB 17 FPGTRCDIQMTQSPSSLSASLGQRVSLTCRASQ 49

## RESULT 7

AAE20204

ID AAE20204 standard; Protein: 146 AA.

XX AC AAE20204;

XX DT 02-FEB-1996 (first entry)

XX DE Anti-idiotypic antibody Idiol7 clone 17KBL.

XX KW Antibody; cancer; CDR; heavy chain; light chain; immunoglobulin;  
XX KW complementarity determining region.

XX OS Mus sp.

FH Key Location/Qualifiers  
 FT Peptide 1..13  
 FT /label= signal\_peptide  
 PN JP07101999-A.  
 PN 18-APR-1995.  
 PD 06-OCT-1993; 93JP-0272950.  
 PF 06-OCT-1993; 93JP-0272950.  
 PR 06-OCT-1993; 93JP-0272950.  
 PX (HAGI/) HAGIWARA Y.  
 PA WPI; 1995-182987/24.  
 PX N-PSDB; AAQ90431.  
 DR Novel anti-idiotype antibody against an human anticancer monoclonal  
 DR antibody - and DNA sequences encoding the antibody, useful in  
 DR pharmacology, medicine and biochemical fields.  
 PX Example 5; Page 19; 28pp; Japanese.  
 PX AAR74960-R74969 are clones of the anti-idiotype antibodies Idio3,  
 CC Idio17, Idio20, Idio27 and Idio33 against a human anticancer monoclonal  
 CC antibody. These antibodies and DNA encoding them are useful in  
 CC pharmacological, medical and biochemical fields of research.  
 XX Sequence 146 AA;  
 SQ  
 Query Match 20.7%; Score 30; DB 16; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-19;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 116 YTEGGGKLEIKRAADAAPTVSIFPPSSKLG 145  
 Db |||||  
 112 YTEGGGKLEIKRAADAAPTVSIFPPSSKLG 141  
 RESULT 8  
 AAY39452  
 ID AAY39452 standard; Protein; 206 AA.  
 AC AAY39452;  
 XX 19-NOV-1999 (first entry)  
 DT Antibody ABX-CBL light chain sequence.  
 DE  
 XX Antibody; CD147; IgM; ABX-CBL; activated T-cell killing; leukaemia;  
 KW activated B-cell; monocyte; graft versus host disease; therapy; cancer;  
 KW organ transplant rejection disease; lymphoma; pancreatic disease;  
 KW autoimmune disease; inflammatory disease; arthritis; binding site.  
 XX Homo sapiens.  
 OS WO9945031-A2.  
 PN 10-SEP-1999.  
 PD 03-MAR-1999; 99WO-US04583.  
 PF 03-MAR-1998; 98US-0034607.  
 PR 03-FEB-1999; 99US-0244253.  
 PX (ABGE-) ABGENIX INC.  
 PA Davis CG, Blacher RW, Corvalan JR, Culwell AR, Green LL, Hales J;  
 PI Havrilla N, Ivanov VE, Lipani JA, Liu Q, Weber RF, Yang X;  
 PX WPI; 1999-540816/45.  
 DR New monoclonal antibody, used for treating e.g. graft versus host  
 PT

PT disease, cancers, autoimmune diseases and inflammatory diseases  
 XX Disclosure; Page 58; 245pp; English.  
 CC This sequence represents the light chain of the antibody ABX-CXL. The  
 CC invention relates to a monoclonal antibody (MAB) with an isotype  
 CC that fixes complement and a variable region that binds to the epitope on  
 CC CD147 bound by the IgM MAB ABX-CBL, providing that the antibody is not  
 CC CBL1. The MAB can selectively kill activated T-cells, activated B-cells  
 CC or resting or activated monocytes. The products and methods can be used  
 CC for treating diseases involving activated T-cells or B-cells or  
 CC monocytes, e.g. graft versus host disease (GVHD), organ transplant  
 CC rejection diseases (e.g. renal transplant, ocular transplant), cancers  
 CC (e.g. cancers of the blood (e.g. leukaemia's and lymphomas) and  
 CC pancreatic), autoimmune diseases (e.g. lupus), and inflammatory diseases  
 CC (e.g. arthritis).  
 SQ Sequence 206 AA;  
 Query Match 20.7%; Score 30; DB 20; Length 206;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-19;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 113 SSPYTFGGGKLEIKRAADAAPTVSIFPPSS 142  
 Db |||||  
 85 SSPYTFGGGKLEIKRAADAAPTVSIFPPSS 114  
 RESULT 9  
 AAW95480  
 ID AAW95480 standard; Protein; 92 AA.  
 XX AAW95480;  
 AC AAW95480;  
 XX 29-MAR-1999 (first entry)  
 DT Mouse derived RT3 phage antibody light chain pattern C genetic sequence.  
 DE Catalytic; antibody; phage display; immunising; phage expression vector;  
 XX prodrug; scfv.  
 KW Mus sp.  
 OS US5855885-A.  
 PN 05-JAN-1999.  
 PD 14-JUL-1994; 94US-0273146.  
 PF 22-JAN-1993; 93US-0007684.  
 PR 14-JUL-1994; 94US-0273146.  
 XX (CHIS/) CHISWELL D.  
 PA (DARS/) DARSLEY M J.  
 PA (FITZ/) FITZGERALD K.  
 PA (KENT/) KENTEN J H.  
 PA (MART/) MARTIN M T.  
 PA (MCCA/) MCCAFFERTY J.  
 PA (SMIT/) SMITH R.  
 PA (TITM/) TITMAS R C.  
 PA (WILL/) WILLIAMS R O.  
 XX Chiswell D, Darsley MJ, Fitzgerald K, Kenten JH;  
 PI Martin MT, McCafferty J, Smith R, Titmas RC, Williams RO;  
 PX WPI; 1999-105036/09.  
 DR N-PSDB; AAX00879.  
 DR Production of catalytic antibodies displayed on bacteriophages -  
 XX comprises generating a gene library of antibody-derived domains  
 PT inserting coding into a phage expression vector and isolating the  
 PT catalytic antibodies  
 XX

PS Example 4; Fig 11; 117pp; English.

XX The invention relates to methods for producing catalytic antibodies

CC displayed on a phage. The method comprises: (a) generating a gene library

CC of antibody-derived domains; (b) inserting coding for the domains into a

CC phage expression vector; and (c) isolating the catalytic antibodies. The

CC phage expression vector incorporates a histidine peptide in tandem with a

CC myc peptide. The catalytic antibodies can be isolated by preparing an

CC antigen; optionally immunising an animal with the antigen; generating a

CC library of VH and VL domains from the immunised animal; cloning the VH

CC and VL domains into a phage expression vector to generate phage display

CC antibodies; selecting the selected phage display antibodies which bind specifically

CC to the antigen; screening the selected phage display antibodies for

CC where the phage expression vector incorporates a histidine peptide in

CC tandem with a myc peptide. The processes are used to produce catalytic

CC antibodies, which can be used for in vivo activation of a prodrug. The

CC present sequence represents a genetic sequence of light chain PCR

CC pattern C from mouse derived RT3 phage antibodies.

XX

SQ Sequence 92 AA;

Query Match 20.0%; Score 29; DB 20; Length 92;

Best Local Similarity 100.0%; Pred. No. 8e-19;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPRFSGSRGSDYSLTISSEDFV 104

|||||

Db 48 SGVPRFSGSRGSDYSLTISSEDFV 76

RESULT 10

AAW95476

ID AAW95476 standard; Protein; 92 AA.

AC

AAW95476;

XX

DT 29-MAR-1999 (first entry)

XX

DE Mouse derived RT3 phage antibody light chain pattern A genetic sequence.

XX

KW Catalytic; antibody; phage display; immunising; phage expression vector;

KW prodrug; scFv.

XX

OS Mus sp.

XX

PN US5855885-A.

XX

PD 05-JAN-1999.

XX

PF 14-JUL-1994; 94US-0273146.

XX

PR 22-JAN-1993; 93US-0007684.

PR

PR 14-JUL-1994; 94US-0273146.

XX

PA (CHIS/) CHISWELL D.

PA (DARS/) DARSLEY M. J.

PA (FITZ/) FITZGERALD K.

PA (KENT/) KENTEN J. H.

PA (MART/) MARTIN M. T.

PA (MCCA/) MCCAFFERTY J.

PA (SMIT/) SMITH R.

PA (TITM/) TITMAS R. C.

PA (WILL/) WILLIAMS R. O.

XX

PI Chiswell D, Darsley MJ, Fitzgerald K, Kenten JH;

PI Martin MT, McCafferty J, Smith R, Titmas RC, Williams RO;

XX

XX WPI; 1999-105036/09.

DR

DR N-PSDB; AAX00875.

XX

XX Production of catalytic antibodies displayed on bacteriophages -

PT comprises generating a gene library of antibody-derived domains

inserting coding into a phage expression vector and isolating the

catalytic antibodies

Example 4; Fig 9A-F; 117pp; English.

The invention relates to methods for producing catalytic antibodies

displayed on a phage. The method comprises: (a) generating a gene library

of antibody-derived domains; (b) inserting coding for the domains into a

phage expression vector; and (c) isolating the catalytic antibodies. The

phage expression vector incorporates a histidine peptide in tandem with a

myc peptide. The catalytic antibodies can be isolated by preparing an

antigen; optionally immunising an animal with the antigen; generating a

library of VH and VL domains from the immunised animal; cloning the VH

and VL domains into a phage expression vector to generate phage display

antibodies; selecting the selected phage display antibodies which bind specifically

to the antigen; screening the selected phage display antibodies for

where the phage expression vector incorporates a histidine peptide in

tandem with a myc peptide. The processes are used to produce catalytic

antibodies, which can be used for in vivo activation of a prodrug. The

present sequence represents a genetic sequence of light chain pattern A

from mouse derived RT3 phage antibodies.

XX

SQ Sequence 92 AA;

Query Match 20.0%; Score 29; DB 20; Length 92;

Best Local Similarity 100.0%; Pred. No. 8e-19;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPRFSGSRGSDYSLTISSEDFV 104

|||||

Db 48 SGVPRFSGSRGSDYSLTISSEDFV 76

RESULT 11

AAW70954

ID AAW70954 standard; Protein; 95 AA.

AC

AAW70954;

XX

DT 14-OCT-1998 (first entry)

XX

DE Light chain of monoclonal antibody RS-255.

XX

KW Pathogenic virus; tropism; mucosa; CDR region; monoclonal antibody;

KW respiratory syncytial virus; RSV; VP6 protein; rota virus; RV;

KW viral infection; inhibit; fusion; protection; transcription;

KW antiviral agent; prophylaxis; diagnosis; infection; contamination.

XX

OS Mus sp.

XX

PN FR2758331-A1.

XX

PD 17-JUL-1998.

XX

PF 14-JAN-1997; 97FR-0000300.

XX

PR 14-JAN-1997; 97FR-0000300.

XX

PA (UYBO-) UNIV BOURGOGNE.

XX

PI Bourgeois C, Kohli E, Pothier P;

XX

XX WPI; 1998-390320/34.

XX

XX New peptide(s) recognising viral epitope with tropism to mucosa -

PT useful for, e.g. diagnosing, preventing and treating viral

PT infection(s)

XX

XX Disclosure; Fig 2; 51pp; French.

XX

XX The present sequence represents the light chain of monoclonal antibody

CC RS-255, which is directed against the respiratory syncytial virus (RSV).



CC The specification describes peptides which recognise, by  
 CC antigen-antibody type reactions, at least 1 epitope of a pathogenic virus  
 CC having tropism for the mucosa. AAW70905-16 and AAW70929-46 are analogous  
 CC to CDR regions of monoclonal antibodies specific for RSV. AAW70917-28 are  
 CC analogous to CDR regions of monoclonal antibodies specific for site III  
 CC or IV of the VP6 protein of rota virus (RV). The peptides can neutralise  
 CC viral infections and may also inhibit fusion between infected and  
 CC uninfected cells or cells and viruses. They provide passive or active  
 CC protection and/or inhibit transcription of the virus, so are useful as  
 CC antiviral agents or for prophylaxis, in human or veterinary medicine. The  
 CC peptides can be labelled and used to diagnose infection or contamination  
 CC by the virus. The peptides are particularly directed against RSV or RS  
 CC but may also be used against papilloma, adeno, entero, polio, influenza  
 CC or immune deficiency viruses.

XX  
 SQ Sequence 95 AA;  
 Query Match 20.0%; Score 29; DB 19; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-19;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPRFSGRSGSDYSLTISSEDFV 104  
 |||||||  
 DB 43 SGVPRFSGRSGSDYSLTISSEDFV 71

RESULT 12  
 AAY90824  
 ID AAY90824 standard; Protein; 107 AA.

XX AC AAY90824;  
 XX  
 DT 29-AUG-2000 (first entry)  
 XX  
 DE 520C9 hybridoma VL domain SEQ ID NO:26.

XX Antigen binding site; immunoglobulin; cancer antigen; immunological;  
 KW antibody; tumour; human; mucin; cancer; cytostatic; hybridoma;  
 KW specific binding assay; affinity purification; drug targeting;  
 KW toxin targeting; imaging; genetic; therapeutic.

XX Homo sapiens.  
 XX  
 PN US6054561-A.  
 XX  
 PD 25-APR-2000.  
 XX  
 PF 07-JUN-1995; 95US-0483749.  
 XX  
 PR 21-MAR-1986; 86US-0842476.  
 PR 08-MAY-1988; 88US-0190778.  
 PR 08-FEB-1984; 84US-0577976.  
 PR 11-JAN-1985; 85US-0690750.  
 PR 11-AUG-1994; 94US-0288981.

XX (CHIR ) CHIRON CORP.

PI Ring DB;

XX WPI: 2000-338508/29.  
 DR N-PSDB; AAA38908.

XX Monoclonal antibody capable of binding to human breast cancer antigen  
 PT useful for affinity purification, drug or toxin targeting, imaging, and  
 PT treating cancer

XX Disclosure; Fig 13; 57pp; English.

XX The present invention describes a monoclonal antibody (MAB) (I) that  
 CC binds to a human breast cancer antigen that is also bound by MAb 454C11  
 CC and 520C9 (produced hybridoma ATCC HB8484 and HB8696, respectively). Also  
 CC described is a hybridoma that produces (I). (I) is useful in specific  
 CC binding assays, affinity purification, drug or toxin targeting, imaging,

CC and genetic or immunological therapeutics for various cancers. The  
 CC present sequence represents a VL domain derived from a 520C9 hybridoma,  
 CC which is used in the exemplification of the present invention.

XX SQ Sequence 107 AA;

Query Match 20.0%; Score 29; DB 21; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 9e-19;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPRFSGRSGSDYSLTISSEDFV 104  
 |||||||  
 DB 56 SGVPRFSGRSGSDYSLTISSEDFV 84

RESULT 13  
 AAR21310  
 ID AAR21310 standard; Protein; 108 AA.

XX AC AAR21310;  
 XX  
 DT 21-MAY-1992 (first entry)  
 XX  
 DE Light chain of M1f clone.

XX Fd; bacteriophage; gene III; filamentous; phagemid; capsid; coat;  
 KW pilus; g3p; binding; adsorption; gene VIII; diverse repertoire;  
 KW specific binding pairs; replicable genetic display package.

OS Synthetic.

XX Key Location/Qualifiers  
 FT Region 25..34  
 FT /label= CDR1  
 FT Region 50..56  
 FT /label= CDR2  
 FT Region 89..96  
 FT /label= CDR3

XX WO9201047-A.  
 XX  
 PD 23-JAN-1992.  
 XX  
 PF 10-JUL-1991; 91WO-GB01134.  
 XX  
 PR 15-MAY-1991; 91GB-0010549.  
 PR 10-JUL-1990; 90GB-0015198.  
 PR 19-OCT-1990; 90GB-0022845.  
 PR 12-NOV-1990; 90GB-0024503.  
 PR 06-MAR-1991; 91GB-0004744.

XX (CMB-) CAMBRIDGE ANTIBODY.  
 PA (MEDI-) MED RES COUNCIL.

XX McCafferty J, Pope AR, Johnson KS, Hoogenboom HRJ, Griffiths AD;  
 PI Jackson RH, Holliger KP, Marks JD, Clackson TP, Chiswell DJ;  
 PI Winter GP, Bonnett TP;

XX WPI; 1992-056862/07.

XX Producing members of specific binding pairs - by expression in  
 PT recombinant host cells with a secreting replicable genetic  
 PT display package.

XX Example 46; Fig 52; 109pp; English.

XX The sequence is the light chain of clone M1f encoding an scFv frag-  
 CC ment specific for both hen and turkey egg lysozyme (HEL and TEL).  
 CC The DNA encoding the chain was amplified from a cDNA library prepd.  
 CC from the spleen of an unimmunised mouse. The corresponding heavy  
 CC chain was also amplified from an existing construct, pSWI-VHDI.3  
 CC (ward et al, 1989). The two fragments were assembled via a linker  
 CC to prepare an scFv construct which was ligated into the fdCAT2

CC vector for expression on the surface of fd bacteriophage. In this  
 CC way, the VL domain was replaced by a library of VL domains to allow  
 CC for selection of a broader range of antibody specificities. Several  
 CC clones were isolated which bound to TEL (the parent antibody D1.3  
 CC binds exclusively to HEL). The sequences of the light chains of  
 CC two of these clones, WF1 and M21 are given in AAR21310 and AAR21311  
 CC respectively. The D1.3 light chain is given in AAR21309.  
 CC See also AAR21260-307, 309-312, AAR22450, AAR22565, AAR22567-81.  
 XX  
 SQ Sequence 108 AA;

Query Match 20.0%; Score 29; DB 13; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-19;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPKFRSGRSGSDYSLTISSEDFV 104  
 |||||  
 DB 56 SGVPKFRSGRSGSDYSLTISSEDFV 84

RESULT 14  
 AAY44587  
 ID AAY44587 standard; Protein; 108 AA.

AC AAY44587;  
 DT 04-APR-2000 (first entry)  
 DE Mouse anti-IL-18 antibody light chain variable region.

XX Anti-Interleukin-18 antibody; IL-18; light chain variable region; VL;  
 KW hybridoma #125-2H; mouse; monoclonal antibody #125-2HmAb; interleukin-18;  
 KW antinflammatory; immunosuppressive; leucocytopenic; antialgic;  
 KW antipyretic; antiallergic; hepatotropic; treatment; autoimmune disease;  
 KW immunopathy; inflammatory disorder; immunoreaction.

XX Mus musculus.

XX EP974600-A2.

XX 26-JAN-2000.

XX 24-JUN-1999; 99EP-0304977.

XX 24-JUN-1998; 98JP-0177580.

XX 12-OCT-1998; 98JP-0289044.

XX 22-DEC-1998; 98JP-0365023.

XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.

XX Nishida Y, Okura T, Tanimoto T, Kurimoto M;

XX WPI; 2000-118341/11.

XX N-PSDB; AAZ49534.

XX New artificially produced peptide for neutralizing biological activity  
 PT of interleukin-18, useful for treating and preventing immunopathies,  
 PT inflammatory disorders and autoimmune diseases -

XX Claim 5; Page 17; 36pp; English.

XX The present protein sequence is the mouse anti-Interleukin-18 antibody  
 CC light chain variable region (VL) encoded by cDNA derived from hybridoma  
 CC #125-2H. The nucleotide sequence is used in the production of recombinant  
 CC monoclonal antibody #125-2HmAb, which is capable of neutralising  
 CC biological activities of interleukin-18. The antibody has antialgic,  
 CC antinflammatory, immunosuppressive, leucocytopenic, antipyretic,  
 CC antiallergic and hepatotropic activity and can be used for prevention and  
 CC treatment of autoimmune diseases, immunopathies and inflammatory  
 CC disorders caused by excessive immunoreaction.

XX Sequence 108 AA;

Query Match 20.0%; Score 29; DB 21; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-19;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPKFRSGRSGSDYSLTISSEDFV 104  
 |||||  
 DB 56 SGVPKFRSGRSGSDYSLTISSEDFV 84

RESULT 15

AAE34366

ID AAE34366 standard; Protein; 130 AA.

XX AAE34366;

DT 14-MAY-2003 (first entry)

XX Escherichia coli light chain variable region.

XX S-surface antigen; Hepatitis B virus; HBV; therapy; infection; virucide;  
 KW hepatotropic.

XX Escherichia coli.

XX Key Location/Qualifiers

FT Region 24..40

FT /note= "CDR1"

FT Region 56..62

FT /note= "CDR2"

FT Region 95..102

FT /note= "CDR3"

XX WO200292819-A1.

XX 21-NOV-2002.

XX 15-MAY-2002; 2002WO-KR00905.

XX 16-MAY-2001; 2001KR-0026634.

XX (YUHA-) YUHAN CORP.

XX Lee JW, Ko IY, Kang HK, Song MY, Song TH, Kim CS;

XX WPI; 2003-140281/13.

XX N-PSDB; AAD52607.

XX New light and heavy chain variable regions of a monoclonal antibody  
 PT against the S-surface antigen of the hepatitis B virus (HBV), useful  
 PT for neutralizing or removing HBV, or for preventing or treating HBV  
 PT infection -

XX Claim 2; Page 16; 20pp; English.

XX The invention relates to light and heavy chain variable regions of a  
 CC monoclonal antibody against S-surface antigen of the hepatitis B virus  
 CC (HBV). The variable regions of the antibodies are useful against HBV  
 CC S-surface antigens, e.g. adr, adw, ayr or ayw, particularly for  
 CC neutralising or removing HBV. They may also be employed to treat or  
 CC prevent HBV infection. The present sequence is Escherichia coli light  
 CC chain variable region.

XX Sequence 130 AA;

Query Match 20.0%; Score 29; DB 24; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-18;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 TFGGGTKLEIKRADAPTSTIFFPSSKLG 145  
 |||||  
 DB 102 TFGGGTKLEIKRADAPTSTIFFPSSKLG 130

Search completed: August 30, 2003, 22:09:54  
Job time : 25.8154 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 22:08:56 ; Search time 15.5705 Seconds  
(without alignments)  
1273.781 Million cell updates/sec

Title: US-08-836-455-2

Perfect score: 145

Sequence: 1 MGAPAQILGFLLLFPGRTRC.....IKRDAAPVTSIFPPSSKLG 145

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 510680 seqs, 136781880 residues

Word size : 0

Total number of hits satisfying chosen parameters: 510680

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145	100.0	145	9	US-08-861-294-2
2	145	100.0	145	12	US-10-367-506-2
3	33	22.8	129	9	US-09-007-093-2
4	29	20.0	108	10	US-09-924-099-1
5	29	20.0	130	8	US-08-779-784-35
6	29	20.0	130	14	US-10-146-305-7
7	29	20.0	149	10	US-09-990-205-2
8	29	20.0	149	15	US-10-153-401-2
9	29	20.0	237	10	US-09-924-099-9
10	29	20.0	243	10	US-09-924-099-10
11	29	20.0	243	10	US-09-887-853-6
12	28	19.3	108	12	US-10-268-883-12
13	28	19.3	112	12	US-10-355-780-1
14	28	19.3	130	12	US-10-268-883-11
15	27	18.6	109	12	US-09-943-906-74

16	17.9	26	17.9	124	9	US-09-802-077-4	Sequence 4, Appl1
17	17.9	26	17.9	124	9	US-09-802-096-4	Sequence 4, Appl1
18	17.9	26	17.9	124	11	US-09-925-179-4	Sequence 4, Appl1
19	17.9	26	17.9	142	10	US-09-840-459-102	Sequence 102, Appl
20	17.9	26	17.9	234	12	US-10-281-479A-24	Sequence 24, Appl
21	17.9	26	17.9	238	10	US-09-903-327A-4	Sequence 4, Appl1
22	17.9	26	17.9	238	15	US-10-216-484-11	Sequence 11, Appl
23	17.9	26	17.9	672	11	US-09-900-766-1	Sequence 1, Appl1
24	17.2	25	17.2	33	9	US-09-861-294-9	Sequence 9, Appl1
25	17.2	25	17.2	33	12	US-10-367-506-9	Sequence 9, Appl1
26	17.2	25	17.2	142	9	US-09-797-481-2	Sequence 2, Appl1
27	17.2	25	17.2	142	9	US-09-844-736-4	Sequence 4, Appl1
28	17.2	25	17.2	142	15	US-10-162-396-4	Sequence 4, Appl1
29	16.6	24	16.6	95	10	US-09-943-906-72	Sequence 72, Appl
30	15.9	23	15.9	23	9	US-09-861-294-5	Sequence 5, Appl1
31	15.9	23	15.9	23	12	US-10-367-506-5	Sequence 5, Appl1
32	14.5	21	14.5	121	11	US-09-929-665-11	Sequence 11, Appl
33	14.5	21	14.5	121	11	US-09-929-546-11	Sequence 11, Appl
34	13.8	20	13.8	269	14	US-10-027-770-2	Sequence 2, Appl1
35	13.8	20	13.8	269	14	US-10-027-770-5	Sequence 5, Appl1
36	13.1	19	13.1	129	9	US-09-839-447A-2	Sequence 2, Appl1
37	13.1	19	13.1	129	15	US-10-153-271-2	Sequence 2, Appl1
38	11.7	17	11.7	19	9	US-09-839-447A-25	Sequence 25, Appl
39	11.7	17	11.7	19	15	US-10-153-271-25	Sequence 25, Appl
40	11.0	16	11.0	16	9	US-09-756-899A-3	Sequence 3, Appl1
41	11.0	16	11.0	23	11	US-09-563-222-130	Sequence 130, Appl
42	11.0	16	11.0	108	9	US-09-329-200A-17	Sequence 17, Appl
43	11.0	16	11.0	109	9	US-09-329-200A-7	Sequence 7, Appl1
44	11.0	16	11.0	109	11	US-09-929-665-21	Sequence 21, Appl
45	11.0	16	11.0	109	11	US-09-929-546-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-09-861-294-2  
; Sequence 2, Application US/09861294  
; Patent No. US20020098190A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TITLE OF INVENTION: TUMORS BEARING HMF AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/09/861,294  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 145  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(20)  
US-09-861-294-2

Query Match 100.0%; Score 145; DB 9; Length 145;  
Best Local Similarity 100.0%; Pred. No. 6.4e-122;  
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGAPAQILGFLLLPPGTRCDIQMTQSPSSLSASLGQRVSLTCRASODIGINLHWLQOEP	60
Db	1	MGAPAQILGFLLLPPGTRCDIQMTQSPSSLSASLGQRVSLTCRASODIGINLHWLQOEP	60
QY	61	DGTIKRLIYATSSILGSGVPRFSGRSGSDYSLTISSEDFVAYCYCLOYASSPYTFGG	120
Db	61	DGTIKRLIYATSSILGSGVPRFSGRSGSDYSLTISSEDFVAYCYCLOYASSPYTFGG	120

Qy 121 GTKLEIKRADAAPTVSIFPPSSKLG 145  
|||||  
Db 121 GTKLEIKRADAAPTVSIFPPSSKLG 145

## RESULT 2

US-10-367-506-2  
; Sequence 2, Application US/10367506  
; Publication No. US20030152575A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TITLE OF INVENTION: TUMORS BEARING HMEG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/10/367,506  
; CURRENT FILING DATE: 2003-02-13  
; PRIOR APPLICATION NUMBER: US/09/861,294  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 145  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(20)  
US-10-367-506-2

Query Match 100.0%; Score 145; DB 12; Length 145;  
Best Local Similarity 100.0%; Pred. No. 6.4e-122; Indels 0; Gaps 0;  
Matches 145; Conservative 0; Mismatches 0;  
Qy 1 MGAPAQILGFLLLFPGTRCDIQMTQSPSSLSASLGQVSLTCSRASODIGINLHWLQOEP 60  
|||||  
Db 1 MGAPAQILGFLLLFPGTRCDIQMTQSPSSLSASLGQVSLTCSRASODIGINLHWLQOEP 60  
Qy 61 DGTIKRLIYATSSLGSGVPRKFRSGSGSDYSLTISSEDFVAYYCLQYASSPYTFGG 120  
|||||  
Db 61 DGTIKRLIYATSSLGSGVPRKFRSGSGSDYSLTISSEDFVAYYCLQYASSPYTFGG 120  
Qy 121 GTKLEIKRADAAPTVSIFPPSSKLG 145  
|||||  
Db 121 GTKLEIKRADAAPTVSIFPPSSKLG 145

## RESULT 3

US-09-007-093-2  
; Sequence 2, Application US/09007093  
; Patent No. US2002025315A1  
; GENERAL INFORMATION:  
; APPLICANT: Anand, Naveen N  
; APPLICANT: Barber, Brian H  
; APPLICANT: Cates, George A  
; APPLICANT: Caterini, Judith E  
; APPLICANT: Klein, Michel H  
; TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF  
; TITLE OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/007,093  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION NUMBER: US 08/483,576  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-765  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 129 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-007-093-2

Query Match 22.8%; Score 33; DB 9; Length 129;  
Best Local Similarity 100.0%; Pred. No. 6.4e-22;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 FPGTRCDIQMTQSPSSLSASLGQVSLTCSRASQ 47  
|||||  
Db 17 FPGTRCDIQMTQSPSSLSASLGQVSLTCSRASQ 49

## RESULT 4

US-09-924-099-1  
; Sequence 1, Application US/09924099  
; Patent No. US20020128450A1  
; GENERAL INFORMATION:  
; APPLICANT: NISHIDA, Yoshihiro  
; APPLICANT: OKURA, Takanori  
; APPLICANT: TANIMOTO, Tadao  
; APPLICANT: KURIMOTO, Masashi  
; TITLE OF INVENTION: PEPTIDE  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/924,099  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 33  
; SEQ ID NO 1  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-924-099-1

Query Match 20.0%; Score 29; DB 10; Length 108;  
Best Local Similarity 100.0%; Pred. No. 2.1e-18;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 76 SGVPRKFRSGSGSDYSLTISSEDFV 104  
|||||  
Db 56 SGVPRKFRSGSGSDYSLTISSEDFV 84

## RESULT 5

US-08-779-784-35  
; Sequence 35, Application US/08779784

```
; Publication No. US20020164325A1
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Moses
; APPLICANT: Miller, David J.
; APPLICANT: Asakura, Kunihiko
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM
; TITLE OF INVENTION: REMYELINATION USING MONOCLONAL AUTOANTIBODIES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,784
; FILING DATE: 07-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/692,084
; FILING DATE: 08-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/236,520
; FILING DATE: 29-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; US-08-779-784-35

Query Match 20.0%; Score 29; DB 8; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPRFSGSRGSDYSLTISSESEDFV 104
Db 78 SGVPRFSGSRGSDYSLTISSESEDFV 106

RESULT 6
US-10-146-305-7
; Sequence 7, Application US/10146305
; Publication No. US20020173035A1
; GENERAL INFORMATION:
; APPLICANT: YUHAN CORPORATION
; TITLE OF INVENTION: A VARIABLE REGION OF THE MONOCLONAL ANTIBODY AGAINST THE HBV
; FILE REFERENCE: OVI1440
; CURRENT APPLICATION NUMBER: US/10/146,305
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: KR 10-2001-26634
; PRIOR FILING DATE: 2001-05-16
```

```
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 7
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-146-305-7

Query Match 20.0%; Score 29; DB 14; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 TFGGGTKLEIKRADAAPTVSIFPPSSKLG 145
Db 102 TFGGGTKLEIKRADAAPTVSIFPPSSKLG 130

RESULT 7
US-09-990-205-2
; Sequence 2, Application US/09990205
; Patent No. US20020150572A1
; GENERAL INFORMATION:
; APPLICANT: FOON, Kenneth A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF PSORIASIS
; FILE REFERENCE: 304142000501
; CURRENT APPLICATION NUMBER: US/09/990,205
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: U.S. 09/192,838
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: U.S. 60/065,774
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Mus Musculus
US-09-990-205-2

Query Match 20.0%; Score 29; DB 10; Length 149;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 TFGGGTKLEIKRADAAPTVSIFPPSSKLG 145
Db 121 TFGGGTKLEIKRADAAPTVSIFPPSSKLG 149

RESULT 8
US-10-153-401-2
; Sequence 2, Application US/10153401
; Publication No. US20030114398A1
; GENERAL INFORMATION:
; APPLICANT: Chatterjee, Malaya
; Foon, Kenneth A.
; Chatterjee, Sunil K.
; TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE
; TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```

RESULT 10
US-09-924-099-10
; Sequence 10, Application US/09924099
; Patent No. US20020128450A1
; GENERAL INFORMATION:
; APPLICANT: NISHIDA, Yoshihiro
; APPLICANT: OKURA, Takanori
; APPLICANT: TANIMOTO, Tadao
; APPLICANT: KURIMOTO, Masashi
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/924, 099
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 10
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially produced peptide in the form of a single chain
; OTHER INFORMATION: chain
; OTHER INFORMATION: variable region fragment (scFv) which neutralizes IL-18
US-09-924-099-10

Query Match 20.0%; Score 29; DB 10; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.2e-18;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPKRFGSRGSDYSLTISSESDFV 104
DB 186 SGVPKRFGSRGSDYSLTISSESDFV 214
|||||
|||||

RESULT 11
US-09-887-853-6
; Sequence 6, Application US/09887853
; Patent No. US20020168375A1
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/887,853
; FILING DATE: 21-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,804
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:

```



```
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-887-853-6
Query Match 20.0%; Score 29; DB 10; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.2e-18;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPKRFSGSRGSDYSLTISSEDFV 104
Db 189 SGVPKRFSGSRGSDYSLTISSEDFV 217

RESULT 12
US-10-268-883-12
; Sequence 12, Application US/10268883
; Publication No. US20030138862A1
; GENERAL INFORMATION:
; APPLICANT: TSO, J. Yun
; APPLICANT: Green, Jennifer Macphate
; TITLE OF INVENTION: Anti-HLA-DR Antibodies and the Methods of Using Thereof
; FILE REFERENCE: 05882.0062.NPUS01
; CURRENT APPLICATION NUMBER: US/10/268,883
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: USSN 60/329,178
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: USSN 60/331,965
; PRIOR FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Mouse
US-10-268-883-12
Query Match 19.3%; Score 28; DB 12; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPKRFSGSRGSDYSLTISSEDF 103
Db 56 SGVPKRFSGSRGSDYSLTISSEDF 83

RESULT 13
US-10-355-780-1
; Sequence 1, Application US/10355780
; Publication No. US20030143224A1
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley
; APPLICANT: Safar, Jiri
; APPLICANT: Williamson, Anthony
; APPLICANT: Burton, Dennis
; TITLE OF INVENTION: Antibodies Specific for Ungulate Prp
; FILE REFERENCE: UCAL-194
; CURRENT APPLICATION NUMBER: US/10/355,780
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US/09/627,218B
; PRIOR FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 1
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthesized peptide
US-10-355-780-1
Query Match 19.3%; Score 28; DB 12; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.7e-17;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPKRFSGSRGSDYSLTISSEDF 103
Db 56 SGVPKRFSGSRGSDYSLTISSEDF 83

RESULT 14
US-10-268-883-11
; Sequence 11, Application US/10268883
; Publication No. US20030138862A1
; GENERAL INFORMATION:
; APPLICANT: TSO, J. Yun
; APPLICANT: Green, Jennifer Macphate
; TITLE OF INVENTION: Anti-HLA-DR Antibodies and the Methods of Using Thereof
; FILE REFERENCE: 05882.0062.NPUS01
; CURRENT APPLICATION NUMBER: US/10/268,883
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: USSN 60/329,178
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: USSN 60/331,965
; PRIOR FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Mouse
US-10-268-883-11
Query Match 19.3%; Score 28; DB 12; Length 130;
Best Local Similarity 100.0%; Pred. No. 1.9e-17;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPKRFSGSRGSDYSLTISSEDF 103
Db 78 SGVPKRFSGSRGSDYSLTISSEDF 105

RESULT 15
US-09-943-906-74
; Sequence 74, Application US/09943906
; Patent No. US20020150571A1
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Williamson, R. Anthony
; APPLICANT: Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE Prp
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/943,906
```

;  
; FILING DATE: 30-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/550,374  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bozicevic, Karl  
; REGISTRATION NUMBER: 28,807  
; REFERENCE/DOCKET NUMBER: 06510/059001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-854-5277  
; TELEFAX: 415-854-0875  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 74:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 109 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 74:  
US-09-943-906-74

Query Match 18.6%; Score 27; DB 10; Length 109;  
Best Local Similarity 100.0%; Pred. No. 1.3e-16;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 76 SGVPRFSGRSGSDYSLTISLSEED 102  
Db 56 SGVPRFSGRSGSDYSLTISLSEED 82

Search completed: August 30, 2003, 22:12:32  
Job time : 16.5705 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 22:07:01 ; Search time 10.2181 Seconds  
(without alignments)  
1364.679 Million cell updates/sec

Title: US-08-836-455-2

Perfect score: 145

Sequence: 1 MGAPQAQLGLLLEFPGR.....IKRADAAPTVSIFPPSKLG 145

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	29	20.0	88	2	PL0261
2	29	20.0	106	2	PL0260
3	29	20.0	106	2	PL0259
4	29	20.0	130	1	KVMSM4
5	28	19.3	91	2	S17622
6	28	19.3	98	2	PH1062
7	28	19.3	108	4	B47271
8	28	19.3	197	2	S29593
9	28	19.3	214	2	S68212
10	27	18.6	225	2	S37484
11	26	17.9	219	2	PC4203
12	26	17.9	219	2	S38865
13	24	16.6	225	2	JL0029
14	22	15.2	128	2	PN0445
15	22	15.2	140	2	PN0446
16	22	15.2	217	2	S42772
17	22	15.2	234	2	S14237
18	20	13.8	106	2	PL0262
19	19	13.1	210	2	A56169
20	19	13.1	219	2	S16112
21	19	13.1	219	2	S52028
22	18	12.4	101	2	B28840
23	18	12.4	117	1	KVMS3B
24	17	11.7	107	2	S40366
25	17	11.7	129	2	S32513
26	16	11.0	94	2	PH1063
27	16	11.0	94	2	E33730
28	16	11.0	97	2	PH1064
29	16	11.0	101	2	C28840

30 16 11.0 103 2 S26332 Ig light chain V r  
31 16 11.0 104 2 S26330 Ig kappa chain V r  
32 16 11.0 104 2 S26329 Ig kappa chain V r  
33 16 11.0 106 2 C33936 Ig kappa chain V r  
34 16 11.0 107 2 PL0269 Ig kappa chain V r  
35 16 11.0 115 2 S60066 Ig kappa chain V r  
36 16 11.0 116 2 A27594 Ig kappa chain pre  
37 16 11.0 117 1 K1HU11 Ig kappa chain pre  
38 16 11.0 117 2 S43528 Ig kappa chain V r  
39 16 11.0 117 2 C21056 Ig kappa chain pre  
40 16 11.0 117 2 B21056 Ig kappa chain pre  
41 16 11.0 117 2 S42263 Ig kappa chain V r  
42 16 11.0 117 2 S21668 Ig kappa chain V r  
43 16 11.0 43 S24206 Ig kappa chain V r  
44 16 11.0 117 2 S41809 Ig kappa chain V r  
45 16 11.0 120 2 S46370 Ig kappa chain V-J

#### ALIGNMENTS

##### RESULT 1

PL0261

Ig kappa chain V region (anti-DNA, DP1SVK) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 09-May-1997

C:Accession: PL0261

R:Shlomchik, M.; Mascelli, M.; Shan, H.; Radic, M.Z.; Pisetsky, D.; Marshak-Rothstein

J. Exp. Med. 171, 265-297, 1990

A:Title: Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somat

A:Reference number: PL0231; MUID:90111618; PMID:2104919

A:Accession: PL0261

A:Molecule type: mRNA

A:Residues: 1-88 <SHL>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:1-5/Region: framework 1

F:6-16/Region: complementarity-determining 1

F:17-31/Region: framework 2

F:32-38/Region: complementarity-determining 2

F:39-70/Region: framework 3

F:71-79/Region: complementarity-determining 3

F:80-88/Region: framework 4

Query Match

Best Local Similarity 20.0%; Score 29; DB 2; Length 88;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPRFSGSRSGSYSLTISSLESEDFV 104

DB 38 SGVPRFSGSRSGSYSLTISSLESEDFV 66

##### RESULT 2

PL0260

Ig kappa chain V region (anti-DNA, DP7VK) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 21-Jan-2000

C:Accession: PL0260

R:Shlomchik, M.; Mascelli, M.; Shan, H.; Radic, M.Z.; Pisetsky, D.; Marshak-Rothstein

J. Exp. Med. 171, 265-297, 1990

A:Title: Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somat

A:Reference number: PL0231; MUID:90111618; PMID:2104919

A:Accession: PL0260

A:Molecule type: mRNA

A:Residues: 1-106 <SHL>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:1-23/Region: framework 1

F:16-90/Domain: immunoglobulin homology <IMM>

F:24-34/Region: complementarity-determining 1

F:35-49/Region: framework 2

F:50-56/Region: complementarity-determining 2

F;57-88/Region: framework 3  
F;89-97/Region: complementarity-determining 3  
F;98-106/Region: framework 4

Query Match 20.0%; Score 29; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 9.7e-21;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPKRFSGSGSDYSLTISSEDFV 104  
|||||  
DB 56 SGVPKRFSGSGSDYSLTISSEDFV 84  
|||||

RESULT 3  
PL0259  
Ig kappa chain V region (anti-DNA, Dp11VK) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 21-Jan-2000  
C;Accession: PL0259  
R;Shlomchik, M.; Mascelli, M.; Shan, H.; Radic, M.Z.; Pisetsky, D.; Marshak-Rothstein, A.  
J. Exp. Med. 171, 265-287, 1990  
A;Title: Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic  
A;Reference number: PL0231; MUID:90111618; PMID:2104919  
A;Accession: PL0259  
A;Molecule type: mRNA  
A;Residues: 1-106 <SHL>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;1-23/Region: framework 1  
F;16-90/Domain: immunoglobulin homology <IMM>  
F;24-34/Region: complementarity-determining 1  
F;35-49/Region: framework 2  
F;50-56/Region: complementarity-determining 2  
F;57-88/Region: framework 3  
F;89-97/Region: complementarity-determining 3  
F;98-106/Region: framework 4

Query Match 20.0%; Score 29; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 9.7e-21;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPKRFSGSGSDYSLTISSEDFV 104  
|||||  
DB 56 SGVPKRFSGSGSDYSLTISSEDFV 84  
|||||

RESULT 4  
KVMSM4  
Ig kappa chain precursor V region (MOPC 41) - mouse  
N;Contains: Ig kappa chain precursor V region VK41  
C;Species: Mus musculus (house mouse)  
C;Date: 24-Apr-1984 #sequence\_revision 24-Apr-1984 #text\_change 28-May-1999  
C;Accession: A93211; A93211; A93815; A94239; A01922; A01923  
R;Seidman, J.G.; Max, E.E.; Leder, P.  
Nature 280, 370-375, 1979  
A;Title: A kappa-immunoglobulin gene is formed by site-specific recombination without fu  
A;Reference number: A93211; MUID:79221900; PMID:111146  
A;Accession: A93211  
A;Molecule type: DNA  
A;Residues: 1-130 <PC41>  
A;Accession: B93211  
A;Molecule type: DNA  
A;Residues: 1-117 <VK41>  
A;Cross-references: GB:V00804; GB:J00566; NID:952127; PID:9575660  
A;Note: The sequences were determined from the differentiated gene MOPC 41 and the germ  
R;Burstein, Y.; Schechter, I.  
Proc. Natl. Acad. Sci. U.S.A. 74, 716-720, 1977  
A;Title: Amino acid sequence of the NH-2-terminal extra piece segments of the precursors  
A;Reference number: A93815; MUID:77148916; PMID:403522  
A;Accession: A93815  
A;Molecule type: protein  
A;Residues: 1-33 <BUR>  
A;Note: Met-3 is apparently used as an alternative initiator in 25% of the chains

R;Gray, W.R.; Dreyer, W.J.; Hood, L.  
Science 155, 465-467, 1967  
A;Title: Mechanism of antibody synthesis: size differences between mouse kappa chains  
A;Reference number: A94239; MUID:67056897; PMID:4162931  
A;Accession: A94239  
A;Molecule type: protein  
A;Residues: 23-49, 'B', '51-53, 'LSB', '57-58, '22', '61-62, 'B2', '65-76, 'B', '78-108, 110-130 <GRA  
A;Experimental source: Bence Jones protein MOPC 41  
C;Genetics:  
A;Introns: 19/1  
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (h  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: alternative initiators; heterotetramer; immunoglobulin  
F;1-22/Domain: signal sequence #status experimental <SIG1>  
F;23-22/Domain: signal sequence #status experimental <SIG2>  
F;23-130/Product: Ig kappa chain V region (MOPC 41) #status experimental <M41>  
F;38-112/Domain: immunoglobulin homology <IMM>  
F;43-110/Disulfide bonds: #status predicted

Query Match 20.0%; Score 29; DB 1; Length 130;  
Best Local Similarity 100.0%; Pred. No. 1.1e-20;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPKRFSGSGSDYSLTISSEDFV 104  
|||||  
DB 78 SGVPKRFSGSGSDYSLTISSEDFV 106  
|||||

RESULT 5  
SL7622  
Ig kappa chain V region - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000  
C;Accession: SL7622  
R;Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G.  
Nature 352, 624-628, 1991  
A;Title: Making antibody fragments using phage display libraries.  
A;Reference number: SL7230; MUID:91326098; PMID:1907718  
A;Accession: SL7622  
A;Status: preliminary  
A;Molecule type: nucleic acid  
A;Residues: 1-91 <CLA>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;8-82/Domain: immunoglobulin homology <IMM>

Query Match 19.3%; Score 28; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 8e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPKRFSGSGSDYSLTISSEDFV 103  
|||||  
DB 48 SGVPKRFSGSGSDYSLTISSEDFV 75  
|||||

RESULT 6  
PH1062  
Ig light chain V region (clone 202.105) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000  
C;Accession: PH1062  
R;Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.  
J. Exp. Med. 176, 761-779, 1992  
A;Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective  
A;Reference number: PH0971; MUID:92381444; PMID:1512540  
A;Accession: PH1062  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-98 <TIL>  
A;Experimental source: B cell, strain [NZB x NZW]F1  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: immunoglobulin

F;16-90/Domain: immunoglobulin homology <IMM>

Query Match 19.3%; Score 28; DB 2; Length 98;  
Best Local Similarity 100.0%; Pred. No. 8.5e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPRFSGSRGSDYSLTISSEDF 103  
|||||  
Db 56 SGVPRFSGSRGSDYSLTISSEDF 83

## RESULT 7

B47271 nitrophenyl phosphonate-specific antibody 48G7 light chain VJ - synthetic (fragment)

C;Species: synthetic  
A;Note: Mus musculus (house mouse) gene engineered and expressed in Escherichia coli

C;Date: 21-Sep-1993 #sequence\_revision 11-Aug-1995 #text\_change 11-Aug-1995  
A;Reference number: A47271; MUID:93165660; PMID:8094556

C;Accession: B47271

R;Lesley, S.A.; Patten, P.A.; Schultz, P.G.

Proc. Natl. Acad. Sci. U.S.A. 90, 1160-1165, 1993

A;Title: A genetic approach to the generation of antibodies with enhanced catalytic acti

A;Reference number: A47271; MUID:93165660; PMID:8094556

C;Accession: B47271

A;Molecule type: DNA; protein

A;Residues: 1-108 <LES>

A;Note: sequence extracted from NCBI backbone (NCBIN:124856, NCBIP:124859)

A;Note: parts of this sequence were determined by protein sequencing

F;24-89/Disulfide bonds: #status predicted

Query Match 19.3%; Score 28; DB 4; Length 108;  
Best Local Similarity 100.0%; Pred. No. 9.2e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPRFSGSRGSDYSLTISSEDF 103  
|||||  
Db 57 SGVPRFSGSRGSDYSLTISSEDF 84

## RESULT 8

S29593 Ig kappa chain (W65) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 11-Jan-2000

C;Accession: S29593

R;Seymour, R.

submitted to the EMBL Data Library, February 1991

A;Reference number: S29593

C;Accession: S29593

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-197 <SEY>

A;Cross-references: EMBL:X57856; NID:g52588; PIDN:CAA40991.1; PID:g52589

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotrimer; immunoglobulin

Query Match 19.3%; Score 28; DB 2; Length 197;  
Best Local Similarity 100.0%; Pred. No. 1.5e-19;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 PYTGGGTGKLEIKRADAAPTIVSIFPPSS 142  
|||||  
Db 115 PYTGGGTGKLEIKRADAAPTIVSIFPPSS 142

## RESULT 9

S68212

Ig kappa chain (Mab03-1) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 29-Jul-1997 #sequence\_revision 29-Aug-1997 #text\_change 11-Jan-2000

C;Accession: S68212

R;Takagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka, T.

FEBS Lett. 375, 273-276, 1995

A;Title: Thermostable peroxidase activity with a recombinant antibody L chain-porphyrin

A;Reference number: S68212; MUID:96085223; PMID:7498516

A;Accession: S68212

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-214 <YAK>

A;Cross-references: EMBL:D29668

C;Superfamily: immunoglobulin V region; immunoglobulin homology

Query Match 19.3%; Score 28; DB 2; Length 214;  
Best Local Similarity 100.0%; Pred. No. 1.6e-19;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 PYTGGGTGKLEIKRADAAPTIVSIFPPSS 142  
|||||  
Db 101 PYTGGGTGKLEIKRADAAPTIVSIFPPSS 128

## RESULT 10

S37484

Ig kappa chain - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 11-Jan-2000

C;Accession: S37484

R;Ducancel, F.F.D.

submitted to the EMBL Data Library, February 1993

A;Reference number: S37483

A;Accession: S37484

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-225 <DUC>

A;Cross-references: EMBL:X70424; NID:g406254; PIDN:CAA49869.1; PID:g406255

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotrimer; immunoglobulin

Query Match 18.6%; Score 27; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.6e-18;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 YTFGGGTGKLEIKRADAAPTIVSIFPPSS 142  
|||||  
Db 107 YTFGGGTGKLEIKRADAAPTIVSIFPPSS 133

## RESULT 11

PC4203

Ig kappa chain (monoclonal antibody MabA34) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 11-Jan-2000

C;Accession: PC4203

R;Kwak, J.W.; Lee, D.I.; Choi, B.K.; Cho, W.K.; Lee, S.H.; Han, M.H.

Gene 173, 257-259, 1996

A;Title: Cloning and characterization of cDNAs coding for heavy and light chains of a

A;Reference number: PC4202; MUID:97082978; PMID:8964510

A;Accession: PC4203

A;Molecule type: mRNA

A;Residues: 1-219 <KWA>

A;Cross-references: GB:U29147; NID:g1594225; PIDN:AAC52821.1; PID:g1594226

C;Comment: This protein is specific for human plasma apolipoprotein A-I of high-densi

C;Superfamily: immunoglobulin V region; immunoglobulin homology

F;1-112/Domain: V region #status predicted <VRG>

F;113-219/Domain: C region #status predicted <CRG>

Query Match 17.9%; Score 26; DB 2; Length 219;  
Best Local Similarity 100.0%; Pred. No. 1.5e-17;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 TFGGGTGTGKLEIKRADAAPTIVSIFPPSS 142  
|||||  
Db 102 TFGGGTGTGKLEIKRADAAPTIVSIFPPSS 127

## RESULT 12

S38865

Ig kappa chain - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Jun-2001  
C:Accession: S38865  
R:Kipp, B.; Becker, W.; Schlaak, M.  
submitted to the EMBL Data Library, November 1993  
A:Description: Combination of a defined specificity and desired isotype by cloning of an  
A:Reference number: S38864  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-219 <KIP>  
A:Cross-references: EMBL:Z27396; NID:q416538; PIDN:CA81787.1; PID:q416539  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match 17.9%; Score 26; DB 2; Length 219;  
Best Local Similarity 100.0%; Pred. No. 1.5e-17; Indels 0; Gaps 0;  
Matches 26; Conservative 0; Mismatches 0

Qy 117 TFGGGTKLEIKRADAAPTSTVFPPSS 142  
|||||  
Db 102 TFGGGTKLEIKRADAAPTSTVFPPSS 127

RESULT 13  
JL0029  
Ig kappa chain precursor (RP93) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 11-Jan-2000  
C:Accession: JL0029  
R:Chien, N.C.; Pollock, R.R.; Desaynard, C.; Scharff, M.D.  
J. Exp. Med. 167, 954-973, 1988  
A:Title: Point mutations cause the somatic diversification of IgM and IgG2a antiphospho  
A:Reference number: JL0029; MUID:88171315; PMID:3127529  
A:Accession: JL0029  
A:Molecule type: mRNA  
A:Residues: 1-225 <CHI>  
A:Experimental source: strain BALB/c, cell line RP93 hybridoma cell  
A:Note: the authors translated the codon CGG for residue 106 as pro, ACC for residue 132  
A:Note: the nucleotide sequence shown is inconsistent with authors' translation because  
ect except for four positions shown above  
C:Comment: The protein is an anti-phosphorylcholine antibody.  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:7-225/Product: Ig kappa chain #status predicted <ILC>  
F:7-106/Domain: V region #status predicted <VAR>  
F:107-119/Domain: J region #status predicted <JIR>  
F:120-225/Domain: C region #status predicted <COR>

Query Match 16.6%; Score 24; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.3e-15;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 TFGGGTKLEIKRADAAPTSTVFPP 140  
|||||  
Db 108 TFGGGTKLEIKRADAAPTSTVFPP 131

RESULT 14  
PN0445  
Ig kappa chain precursor V-I region - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000  
C:Accession: PN0445  
R:Kaluza, B.; Betzel, G.; Shao, H.; Diamantsehn, T.; Weidle, U.H.  
Gene 122, 321-328, 1992  
A:Title: A general method for chimerization of monoclonal antibodies by inverse polymera  
A:Reference number: PN0444; MUID:93138402; PMID:1339379  
A:Accession: PN0445  
A:Molecule type: mRNA  
A:Residues: 1-128 <KAL>  
A:Cross-references: GB:L02347

C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:1-10/Domain: signal sequence #status predicted <SIG>  
F:11-128/Product: Ig light chain kappa-1 V region #status predicted <MAT>  
F:26-100/Domain: immunoglobulin homology <IMM>

Query Match 15.2%; Score 22; DB 2; Length 128;  
Best Local Similarity 100.0%; Pred. No. 7.3e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 TFGGGTKLEIKRADAAPTSTVF 138  
|||||  
Db 107 TFGGGTKLEIKRADAAPTSTVF 128

RESULT 15  
PN0446  
Ig kappa chain precursor V-II region - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000  
C:Accession: PN0446  
R:Kaluza, B.; Betzel, G.; Shao, H.; Diamantsehn, T.; Weidle, U.H.  
Gene 122, 321-328, 1992  
A:Title: A general method for chimerization of monoclonal antibodies by inverse polym  
A:Reference number: PN0444; MUID:93138402; PMID:1339379  
A:Accession: PN0446  
A:Molecule type: mRNA  
A:Residues: 1-140 <KAL>  
A:Cross-references: GB:L02345  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-140/Product: Ig light chain kappa-2 V region #status predicted <MAT>  
F:36-114/Domain: immunoglobulin homology <IMM>

Query Match 15.2%; Score 22; DB 2; Length 140;  
Best Local Similarity 100.0%; Pred. No. 7.9e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 TFGGGTKLEIKRADAAPTSTVF 138  
|||||  
Db 119 TFGGGTKLEIKRADAAPTSTVF 140

Search completed: August 30, 2003, 22:11:49  
Job time : 10.2181 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 22:04:41 ; Search time 8.75839 Seconds  
(without alignments)  
778.553 Million cell updates/sec

Title: US-08-836-455-2  
Perfect score: 145  
Sequence: 1 MGAPAQILGFLLLFPQTRC.....IKRADAPTIVSIIPPSSKLG 145

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	29	20.0	130	1 KV5G_MOUSE	P01639 mus musculus
2	18	12.4	117	1 KV5H_MOUSE	P01641 mus musculus
3	16	11.0	117	1 KV1L_HUMAN	P01601 homo sapien
4	16	11.0	129	1 KV1W_HUMAN	P04431 homo sapien
5	16	11.0	129	1 KV1X_HUMAN	P04432 homo sapien
6	14	9.7	106	1 KAC_MOUSE	P01837 mus musculus
7	14	9.7	108	1 KV1A_HUMAN	P01593 homo sapien
8	14	9.7	108	1 KV1B_HUMAN	P01594 homo sapien
9	14	9.7	108	1 KV1G_HUMAN	P01599 homo sapien
10	14	9.7	108	1 KV1H_HUMAN	P01600 homo sapien
11	14	9.7	108	1 KV1O_HUMAN	P01607 homo sapien
12	14	9.7	108	1 KV1P_HUMAN	P01608 homo sapien
13	14	9.7	108	1 KV1Q_HUMAN	P01609 homo sapien
14	14	9.7	108	1 KV1R_HUMAN	P01610 homo sapien
15	14	9.7	108	1 KV1Y_HUMAN	P08362 homo sapien
16	14	9.7	108	1 KV5Q_MOUSE	P01650 mus musculus
17	14	9.7	108	1 KV5S_MOUSE	P01652 mus musculus
18	14	9.7	108	1 KV5T_MOUSE	P01653 mus musculus
19	14	9.7	113	1 KV2E_MOUSE	P03976 mus musculus
20	14	9.7	113	1 KV2F_MOUSE	P01630 mus musculus
21	13	9.0	106	1 KACA_RAT	P01836 rattus norv
22	13	9.0	106	1 KACB_RAT	P01835 rattus norv
23	13	9.0	111	1 KV3D_MOUSE	P03977 mus musculus
24	13	9.0	111	1 KV3H_MOUSE	P01660 mus musculus
25	13	9.0	111	1 KV3J_MOUSE	P01662 mus musculus
26	13	9.0	111	1 KV3L_MOUSE	P01664 mus musculus
27	13	9.0	112	1 KV1U_HUMAN	P01613 homo sapien
28	13	9.0	112	1 KV2D_MOUSE	P01629 mus musculus
29	13	9.0	112	1 KV3B_MOUSE	P01655 mus musculus
30	13	9.0	136	1 KV5B_MOUSE	P01634 mus musculus
31	12	8.3	107	1 KV6A_MOUSE	P01675 mus musculus
32	12	8.3	107	1 KV6D_MOUSE	P01678 mus musculus
33	12	8.3	108	1 KV1M_HUMAN	P01605 homo sapien

ALIGNMENTS

RESULT 1

ID	KV5G_MOUSE	STANDARD;	PRT;	130 AA.
AC	P01639; P01640;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DE	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Ig kappa chain V-V region MOPC 41 precursor			
OS	Mus musculus (Mouse)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=79221900; PubMed=111146;			
RA	Seldman J.G., Max E.E., Leder P.;			
RT	"A kappa-immunoglobulin gene is formed by site-specific recombination			
RT	without further somatic mutation.";			
RL	Nature 280:370-375(1979).			
RN	[2]			
RP	SEQUENCE OF 1-33.			
RX	MEDLINE=77148916; PubMed=403522;			
RA	Burstein Y., Schechter I.;			
RT	"Amino acid sequence of the NH2-terminal extra piece segments of the			
RT	precursors of mouse immunoglobulin lambda1-type and kappa-type light			
RL	chains.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 74:716-720(1977).			
RP	SEQUENCE OF 23-130.			
RX	MEDLINE=67056897; PubMed=4162931;			
RA	Gray W.R., Dreyer W.J., Hood L.;			
RT	"Mechanism of antibody synthesis: size differences between mouse			
RT	kappa chains.";			
RL	Science 155:465-467(1967).			
CC	-1- MISCELLANEOUS: THIS PRECURSOR WAS SYNTHESIZED IN A CELL-FREE			
CC	SYSTEM DIRECTED BY MRNA ISOLATED FROM MYELOMA POLYSOMES.			
CC	-1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.			
DR	PIR: A93211; KVM5M4.			
DR	HSP; P01607; IREI.			
DR	InterPro: IPR007110; Ig-like.			
DR	InterPro: IPR003006; Ig_MHC.			
DR	InterPro: IPR003596; Ig_V.			
DR	Pfam: PF00047; Ig; 1.			
DR	SMART; SM00406; IGV; 1.			
KW	PROSITE; PS50835; IG_LIKE; 1.			
FT	IMMUNOGLOBULIN V region; Signal; Bence-Jones protein.			
FT	SIGNAL 1 22			
FT	CHAIN 23 130			
FT	IG KAPPA CHAIN V-V REGION MOPC 41.			
FT	DOMAIN 23 45			
FT	FRAMEWORK-1.			
FT	COMPLEMENTARITY-DETERMINING-1.			
FT	DOMAIN 46 56			
FT	FRAMEWORK-2.			
FT	DOMAIN 57 71			
FT	COMPLEMENTARITY-DETERMINING-2.			
FT	DOMAIN 72 78			
FT	FRAMEWORK-3.			
FT	DOMAIN 79 110			
FT	COMPLEMENTARITY-DETERMINING-3.			
FT	DOMAIN 111 119			
FT	FRAMEWORK-4.			
FT	DOMAIN 120 129			
FT	BY SIMILARITY.			
FT	DISULFID 45 110			

34	12	8.3	108	1	KV5J_MOUSE	P01643 mus musculus
35	12	8.3	108	1	KV5K_MOUSE	P01644 mus musculus
36	12	8.3	108	1	KV5L_MOUSE	P01645 mus musculus
37	12	8.3	108	1	KV5M_MOUSE	P01646 mus musculus
38	12	8.3	108	1	KV5N_MOUSE	P01647 mus musculus
39	12	8.3	108	1	KV5O_MOUSE	P01648 mus musculus
40	12	8.3	108	1	KV5P_MOUSE	P01649 mus musculus
41	12	8.3	109	1	KV1T_HUMAN	P01612 homo sapien
42	12	8.3	113	1	KV2G_MOUSE	P01631 mus musculus
43	12	8.3	149	1	KV5A_MOUSE	P01633 mus musculus
44	11	7.6	108	1	KV1E_HUMAN	P01597 homo sapien
45	11	7.6	108	1	KV5R_MOUSE	P01651 mus musculus

```

FT VARIANT      1      2      MISSING (IN 25% OF THE MOLECULES).
FT NON_TER     130     130
SQ SEQUENCE     130 AA: 14311 MW: 5EPE0FE71D5F1BEC CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 1; Length 130;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGVPRFSGSRSGDYSLTISSEDFV 104
DB 78 SGVPRFSGSRSGDYSLTISSEDFV 106

RESULT 2
KV5H_MOUSE
ID KV5H_MOUSE STANDARD; PRT; 117 AA.
AC P01641;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region MOPC 173B precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81064681; PubMed=6777049;
RA Max E.E., Seidman J.G., Miller H., Leder P.;
RT "Variation in the crossover point of kappa immunoglobulin gene V-J
recombination: evidence from a cryptic gene.";
RL Cell 21:793-799(1980).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; K00880; AAA39031.1; -
DR PIR; A01924; KVM3B.
DR HSP; P01607; IRE1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL      1      22
FT CHAIN       23     117 IG KAPPA CHAIN V-V REGION MOPC 173B.
FT DOMAIN      23     45 FRAMEWORK-1.
FT DOMAIN      46     56 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN      57     71 FRAMEWORK-2.
FT DOMAIN      72     78 FRAMEWORK-3.
FT DOMAIN      79     110 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN      111    >117 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID    45     110 BY SIMILARITY.
FT NON_TER    117     117
SQ SEQUENCE    117 AA: 12954 MW: 24B3D4B9AC2E4D6C CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 117;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 RCDIQMTQSPSSLSASLG 36
DB 21 RCDIQMTQSPSSLSASLG 38

RESULT 3
KV11_HUMAN
ID KV11_HUMAN STANDARD; PRT; 117 AA.

```

```

AC P01601;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region HK101 precursor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81098966; PubMed=6779204;
RA Bentley D.L., Rabbitts T.H.;
RT "Human immunoglobulin variable region genes -- DNA sequences of two V
kappa genes and a pseudogene.";
RL Nature 288:730-733(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83129397; PubMed=6402305;
RA Bentley D.L., Rabbitts T.H.;
RT "Evolution of immunoglobulin V genes: evidence indicating that
recently duplicated human V kappa sequences have diverged by gene
conversion.";
RL Cell 32:181-189(1983).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; K01322; AAA58930.1; -
DR EMBL; K01324; AAA58932.1; -
DR EMBL; V00558; CAA23824.1; -
DR PIR; A01881; K1HU11.
DR HSP; P01607; IRE1.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL      1      22
FT CHAIN       23    >117 IG KAPPA CHAIN V-I REGION HK101.
FT DOMAIN      23     45 FRAMEWORK-1.
FT DOMAIN      46     56 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN      57     71 FRAMEWORK-2.
FT DOMAIN      72     78 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN      79     110 FRAMEWORK-3.
FT DOMAIN      111    >117 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID    45     110 BY SIMILARITY.
FT NON_TER    117     117
SQ SEQUENCE    117 AA: 12799 MW: D7D0FF3718CEF587 CRC64;

Query Match
Best Local Similarity 100.0%; Score 16; DB 1; Length 117;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 RCDIQMTQSPSSLSAS 34
DB 21 RCDIQMTQSPSSLSAS 36

RESULT 4
KV1W_HUMAN
ID KV1W_HUMAN STANDARD; PRT; 129 AA.
AC P04431;

```



```

DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region Walker precursor.
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA MEDLINE=85014148; PubMed=6091049;
RA Klobbeck H.G., Combriato G., Zachau H.G.;
RT "Immunoglobulin genes of the kappa light chain type from two human
RL lymphoid cell lines are closely related.";
RL Nucleic Acids Res. 12:6995-7006(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X00965; CAA25477.1; ALT_TERM.
DR HSRP; P01607; IREI.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 22
FT CHAIN 23 129 IG KAPPA CHAIN V-I REGION WALKER.
FT DOMAIN 23 45 FRAMEWORK-1.
FT DOMAIN 46 56 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 57 71 FRAMEWORK-2.
FT DOMAIN 72 78 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 79 110 FRAMEWORK-3.
FT DOMAIN 111 119 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 120 129 FRAMEWORK-4.
FT DISULFID 45 110 BY SIMILARITY.
FT NON_TER 129 129
SQ SEQUENCE 129 AA; 14069 MW; F941FA07D4AFC2F9 CRC64;

Query Match 11.0%; Score 16; DB 1; Length 129;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 RCDIQMTQSPSSLSAS 34
Db 21 RCDIQMTQSPSSLSAS 36
|||||
RESULT 5
KVIX_HUMAN STANDARD; PRT; 129 AA.
AC P04432;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region Daudi precursor.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.

```

```

RX MEDLINE=85014148; PubMed=6091049;
RA Klobbeck H.G., Combriato G., Zachau H.G.;
RT "Immunoglobulin genes of the kappa light chain type from two human
RL lymphoid cell lines are closely related.";
RL Nucleic Acids Res. 12:6995-7006(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X00966; CAA25478.1; ALT_TERM.
DR HSRP; P80362; IWL.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 22
FT CHAIN 23 129 IG KAPPA CHAIN V-I REGION DAUDI.
FT DOMAIN 23 45 FRAMEWORK-1.
FT DOMAIN 46 56 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 57 71 FRAMEWORK-2.
FT DOMAIN 72 78 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 79 110 FRAMEWORK-3.
FT DOMAIN 111 119 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 120 129 FRAMEWORK-4.
FT DISULFID 45 110 BY SIMILARITY.
FT NON_TER 129 129
SQ SEQUENCE 129 AA; 14235 MW; CAF076BC7E5574C8 CRC64;

Query Match 11.0%; Score 16; DB 1; Length 129;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 RCDIQMTQSPSSLSAS 34
Db 21 RCDIQMTQSPSSLSAS 36
|||||
RESULT 6
KAC_MOUSE STANDARD; PRT; 106 AA.
AC P01837;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig kappa chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE (MOPC 21).
RP MEDLINE=73053310; PubMed=4638343;
RA Svasti J., Milstein C.;
RT "The complete amino acid sequence of a mouse kappa light chain.";
RL Biochem. J. 128:427-444(1972).
RN [2]
RP DISULFIDE BONDS (MOPC 21).
RX MEDLINE=73008889; PubMed=5073237;
RA Svasti J., Milstein C.;
RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
RL Biochem. J. 126:837-850(1972).

```

RN RP SEQUENCE FROM N.A., AND REVISIONS TO 53-59.  
RX MEDLINE=79084137; PubMed=103625;  
RA Hamlyn P.H., Brownlee G.G., Cheng C.-C., Gait M.J., Milstein C.;  
RT "Complete sequence of constant and 3' noncoding regions of an  
FT immunoglobulin mRNA using the dideoxynucleotide method of RNA  
RL sequencing.";  
RL Cell 15:1067-1075(1978).  
[4]  
RN RP SEQUENCE FROM N.A.  
RX MEDLINE=82059477; PubMed=6170937;  
RA Hamlyn P.H., Gait M.J., Milstein C.;  
RT "Complete sequence of an immunoglobulin mRNA using specific priming  
FT and the dideoxynucleotide method of RNA sequencing.";  
RL Nucleic Acids Res. 9:4485-4494(1981).  
[5]  
RN RP SEQUENCE FROM N.A.  
RX MEDLINE=81191915; PubMed=6262318;  
RA Max E.E., Maizel J.V. Jr., Leder P.;  
RT "The nucleotide sequence of a 5.5-kilobase DNA segment containing the  
FT mouse kappa immunoglobulin J and C region genes.";  
RL J. Biol. Chem. 256:5116-5120(1981).  
[6]  
RN RP SEQUENCE FROM N.A.  
RX MEDLINE=81198949; PubMed=6785724;  
RA Altenburger W., Neumayer P.S., Steinmetz M., Zachau H.G.;  
RT "DNA sequence of the constant gene region of the mouse immunoglobulin  
FT kappa chain.";  
RL Nucleic Acids Res. 9:971-981(1981).  
[7]  
RN RP SEQUENCE FROM N.A.  
RX MEDLINE=88329081; PubMed=3138116;  
RA de Waele P., Feys V., van de Voorde A., Molemans F., Fiers W.;  
RT "Expression in non-lymphoid cells of mouse recombinant immunoglobulin  
FT directed against the tumour marker human placental alkaline  
RT phosphatase.";  
RL Eur. J. Biochem. 176:287-295(1988).  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/announcements/](http://www.isb-sib.ch/announcements)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; V00807; CAA24189.1; --  
DR PIR; B90262; KIMS.  
DR PDB; 1AIF; 01-FEB-97.  
DR PDB; 1FSK; 02-OCT-00.  
DR PDB; 1IGC; 03-JUN-95.  
DR PDB; 1KB5; 08-APR-98.  
DR PDB; 1KCR; 11-MAY-02.  
DR PDB; 1KCS; 11-MAY-02.  
DR PDB; 1KCU; 11-MAY-02.  
DR PDB; 1KCV; 11-MAY-02.  
DR PDB; 25C8; 09-JUL-99.  
DR MGD; MGI:96495; Igk-C.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig\_ch.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00407; IGcl; 1.  
DR PROSITE; PS0835; IG\_LIKE; 1.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region; 3D-structure.  
FT NON\_TER 1 1 IG-LIKE.  
FT DOMAIN 5 102  
FT DISULFID 26 86  
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
FT SEQUENCE 106 AA; 11778 MW; 4B51FF5EF49BAEB5 CRC64;

Query Match 9.7%; Score 14; DB 1; Length 106;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 129 ADAAPTVSIFPPSS 142  
Db 1 ADAAPTVSIFPPSS 14  
RESULT 7  
KV1A\_HUMAN STANDARD; PRT; 108 AA.  
ID KV1A\_HUMAN  
AC P01593;  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DE 21-JUL-1986 (Rel. 01, Last sequence update)  
DE 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-I region AG.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
CX NCBI\_TaxID=9606;  
[1]  
RN RP SEQUENCE  
RX MEDLINE=69234734; PubMed=4893682;  
RA Titani K., Shinoda T., Putnam F.W.;  
RT "The amino acid sequence of a kappa type Bence-Jones protein. 3. The  
FT complete sequence and the location of the disulfide bridges.";  
RL J. Biol. Chem. 244:3550-3560(1969).  
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.  
CC -1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.  
DR PIR; A01861; KIHUAG.  
DR HSP; P01607; IREI.  
DR GO; GO:0005576; C:extracellular; NAS.  
DR GO; GO:0003823; F:antigen binding activity; NAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGv; 1.  
DR PROSITE; PS0835; IG\_LIKE; 1.  
KW Immunoglobulin V region; Bence-Jones protein.  
FT DOMAIN 1 23 FRAMEWORK-1.  
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 35 49 FRAMEWORK-2.  
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 57 88 FRAMEWORK-3.  
FT DOMAIN 89 97 FRAMEWORK-4.  
FT DOMAIN 98 107  
FT DISULFID 23 88  
FT NON\_TER 108 108  
FT SEQUENCE 108 AA; 11992 MW; E3B3B246C18F0C4F CRC64;  
Query Match 9.7%; Score 14; DB 1; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 21 DIQWTQSPSLSAS 34  
Db 1 DIQWTQSPSLSAS 14  
RESULT 8  
KV1B\_HUMAN STANDARD; PRT; 108 AA.  
ID KV1B\_HUMAN  
AC P01594;  
DT 21-JUL-1986 (Rel. 01, Created)  
DE 21-JUL-1986 (Rel. 01, Last sequence update)  
DE 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ig kappa chain V-I region AU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.



```

FT  DOMAIN      50 56 COMPLEMENTARITY-DETERMINING-2.
FT  DOMAIN      57 88 FRAMEWORK-3.
FT  DOMAIN      89 97 COMPLEMENTARITY-DETERMINING-3.
FT  DOMAIN      98 107 FRAMEWORK-4.
FT  DISULFID     23 88 BY SIMILARITY.
FT  NON_TER     108 108
SQ  SEQUENCE 108 AA; 11671 MW; 08D3A6160D8D0618 CRC64;

Query Match          9.7%; Score 14; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DIQMTQSPSSLSAS 34
    | | | | | | | | | |
DB 1 DIQMTQSPSSLSAS 14

RESULT 11
KV10_HUMAN
ID KV10_HUMAN STANDARD; PRT; 108 AA.
AC P01607;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig kappa chain V-1 region Rel.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76023758; PubMed=809329;
RA Palm W., Hilschmann N.;
RT "The primary structure of a crystalline monoclonal immunoglobulin
RT kappa-type L-chain, subgroup I (Bence-Jones protein Rel); isolation
RT and characterization of the tryptic peptides; the complete amino acid
RT sequence of the protein; a contribution to the elucidation of the
RT three-dimensional structure of antibodies, in particular their
RT combining site."
RL Hoppe-Seyler's Z. Physiol. Chem. 356:167-191(1975).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=76039968; PubMed=1182131;
RA Epp O., Lattman E.E., Schiffer M., Huber R., Palm W.;
RT "The molecular structure of a dimer composed of the variable portions
RT of the Bence-Jones protein REI refined at 2.0-A resolution."
RL Biochemistry 14:4943-4952(1975).
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (1,2)
CC MARKER.
CC -1- MISCELLANEOUS: THIS IS A BENICE-JONES PROTEIN.
DR PIR; A91663; KIHURE.
DR PDB; IRE1; 17-FEB-84.
DR PDB; IAR2; 12-NOV-97.
DR PDB; 1BWV; 29-DEC-99.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Bence-Jones protein; 3D-structure.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 35 49 FRAMEWORK-2.
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 23 88 FRAMEWORK-4.
FT NON_TER 98 107
FT SEQUENCE 108 AA; 11671 MW; 08D3A6160D8D0618 CRC64;

```

```

FT  STRAND      10 13
FT  TURN       15 16
FT  STRAND     19 25
FT  TURN       30 31
FT  STRAND     33 38
FT  TURN       40 41
FT  STRAND     44 49
FT  TURN       53 54
FT  STRAND     56 57
FT  TURN       60 61
FT  STRAND     62 67
FT  TURN       68 69
FT  STRAND     70 75
FT  HELIX      80 82
FT  STRAND     85 90
FT  STRAND     97 98
FT  STRAND    102 106
FT  NON_TER    108 108
SQ  SEQUENCE 108 AA; 11902 MW; 9E8143E1188BCE2A CRC64;

Query Match          9.7%; Score 14; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DIQMTQSPSSLSAS 34
    | | | | | | | | | |
DB 1 DIQMTQSPSSLSAS 14

RESULT 12
KV1P_HUMAN
ID KV1P_HUMAN STANDARD; PRT; 108 AA.
AC P01608;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-1 region Roy.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=68362076; PubMed=5595110;
RA Hilschmann N.;
RT "Chemical structure of 2 kappa-type Bence Jones proteins (Roy and
RT Cum.)."
RL Hoppe-Seyler's Z. Physiol. Chem. 348:1077-1080(1967).
RN [2]
RP REVISIONS TO 39 AND 41.
RA Hilschmann N., Barnikol H.U., Hess M., Langer B., Ponstingl H.,
RA Steinmetz-Kayne M., Suter L., Watanabe S.;
RL (In) Franke F., Shugar D. (eds.);
RL Gamma globulins: structure and function, pp.57-74, Academic Press,
RL New York (1969).
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (1,2)
CC MARKER.
CC -1- MISCELLANEOUS: THIS IS A BENICE-JONES PROTEIN.
DR PIR; A91638; KIHURY.
DR HSSP; P80362; 1WTL.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Bence-Jones protein.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.

```

FT DOMAIN 35 49 FRAMEWORK-2.  
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 57 88 FRAMEWORK-3.  
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.  
FT DOMAIN 98 107 FRAMEWORK-4.  
FT DISULFID 23 88 BY SIMILARITY.  
FT NON\_TER 108 108  
SQ SEQUENCE 108 AA; 11782 MW; F5ACEDESA313DF3A CRC64;

Query Match 9.7%; Score 14; DB 1; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DIQQTSPSSLSAS 34  
Db 1 DIQQTSPSSLSAS 14

## RESULT 13

KV1Q\_HUMAN STANDARD; PRT; 108 AA.  
AC P01609;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-I region Sw.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RA MEDLINE=75059271; PubMed=4435756;  
RX Eulitz M., Hilschmann N.;  
RT "The primary structure of a human immunoglobulin L-chain of  
kappa-type (Bence-Jones protein Scw.), II: The chymotryptic peptides  
and the complete amino acid sequence.",  
RL Hoppe-Seyler's Z. Physiol. Chem. 355:842-866(1974).  
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (1,2)  
MARKER.  
CC -1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.  
DR PIR: A01875; KIHUS.  
DR HSP; P01607; IREI.  
DR GO: GO:0005576; C:extracellular; NAS.  
DR GO: GO:0003823; F:antigen binding activity; NAS.  
DR GO: GO:0006955; P:immune response; NAS.  
DR InterPro: IPR007110; Ig-like.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS0835; IG\_LIKE; 1.  
KW Immunoglobulin V region; Bence-Jones protein.  
FT DOMAIN 1 23 FRAMEWORK-1.  
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 35 49 FRAMEWORK-2.  
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 57 88 FRAMEWORK-3.  
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.  
FT DOMAIN 98 107 FRAMEWORK-4.  
FT DISULFID 23 88 BY SIMILARITY.  
FT NON\_TER 108 108  
SQ SEQUENCE 108 AA; 11764 MW; 32CECD9F644414 CRC64;

Query Match 9.7%; Score 14; DB 1; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DIQQTSPSSLSAS 34  
Db 1 DIQQTSPSSLSAS 14

## RESULT 14

KV1R\_HUMAN STANDARD; PRT; 108 AA.  
AC P01610;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-I region WEA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=83273707; PubMed=6410398;  
RA Goni F., Frangione B.;  
RT "Amino acid sequence of the Fv region of a human monoclonal IgM  
(protein WEA) with antibody activity against 3,4-pyruvylated  
galactose in Klebsiella polysaccharides K30 and K33.",  
RL Proc. Natl. Acad. Sci. U.S.A. 80:4837-4841(1983).  
CC -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM A MONOCLONAL ANTIBODY  
AGAINST 3,4-PYRUVYLATED GALACTOSE AND ISOLATED FROM A PATIENT WITH  
WALDENSTROM'S MACROGLOBULINEMIA.  
DR PIR: A01876; KIHWE.  
DR HSP; P80362; IWTL.  
DR GO: GO:0005576; C:extracellular; NAS.  
DR GO: GO:0003823; F:antigen binding activity; NAS.  
DR GO: GO:0006955; P:immune response; NAS.  
DR InterPro: IPR007110; Ig-like.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS0835; IG\_LIKE; 1.  
KW Immunoglobulin V region; Monoclonal antibody.  
FT DOMAIN 1 23 FRAMEWORK-1.  
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 35 49 FRAMEWORK-2.  
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 57 88 FRAMEWORK-3.  
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.  
FT DOMAIN 98 107 FRAMEWORK-4.  
FT DISULFID 23 88 BY SIMILARITY.  
FT NON\_TER 108 108  
SQ SEQUENCE 108 AA; 11840 MW; 9249B61F0945618C CRC64;

Query Match 9.7%; Score 14; DB 1; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DIQQTSPSSLSAS 34  
Db 1 DIQQTSPSSLSAS 14

## RESULT 15

KV1Y\_HUMAN STANDARD; PRT; 108 AA.  
AC P80362;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ig kappa chain V-I region WAT.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE, AND X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RX MEDLINE=95086080; PubMed=7993911;  
RA Huang D.-B., Chang C.-H., Ainsworth C., Bruenger A.T., Eulitz M.,  
RA Solomon A., Stevens F.J., Schiffer M.;  
RT "Comparison of crystal structures of two homologous proteins:

Query Match 9.7%; Score 14; DB 1; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 DIQQTSPSSLSAS 34  
Db 1 DIQQTSPSSLSAS 14

RT structural origin of altered domain interactions in immunoglobulin  
RL light-chain dimers.";  
RN Biochemistry 33:14848-14857(1994).  
RP [2]

RX SEQUENCE OF 1-35.  
RA MEDLINE=81267384; PubMed=6167731;  
RA Stevens F.J., Westholm F.A., Panagiotopoulos N., Schiffer M.,  
Popp R.A., Solomon A.;  
RT "Characterization and preliminary crystallographic data on the VL-  
RT related fragment of the human kappa Bence Jones protein Wat.";  
RL J. Mol. Biol. 147:185-193(1981).  
CC -1- MISCELLANEOUS: THIS IS A BENICE-JONES PROTEIN.  
DR PDB: 1WTL; 01-NOV-94.  
DR GO: GO:0005576; C:extracellular; NAS.  
DR GO: GO:0003823; F:antigen binding activity; NAS.  
DR GO: GO:0006955; P:immune response; NAS.  
DR InterPro: IPR007110; Ig-like.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00406; IGV; 1.  
DR PROSITE: PS00835; IG\_LIKE; 1.  
KW Immunoglobulin V region; Bence-Jones protein; 3D-structure.  
FT DOMAIN 1 23 FRAMEWORK-1.  
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 35 49 FRAMEWORK-2.  
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 57 88 FRAMEWORK-3.  
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.  
FT DOMAIN 98 107 FRAMEWORK-4.  
FT DISULFID 23 88 BY SIMILARITY.  
FT CONFLICT 30 31 TN -> SD (IN REF. 2).  
FT STRAND 4 7  
FT STRAND 10 13  
FT TURN 15 16  
FT STRAND 19 25  
FT TURN 30 31  
FT STRAND 33 38  
FT TURN 40 41  
FT STRAND 45 49  
FT TURN 50 52  
FT STRAND 53 54  
FT TURN 56 57  
FT TURN 60 61  
FT STRAND 62 67  
FT TURN 68 69  
FT STRAND 70 75  
FT HELIX 80 82  
FT STRAND 84 90  
FT STRAND 98 98  
FT STRAND 102 106  
FT NON\_TER 108 108  
SQ SEQUENCE 108 AA; 11737 MW; D9D941B3F0FAE697 CRC64;

Query Match 9.78; Score 14; DB 1; Length 108;  
Best Local Similarity 100.08; Pred. No. 1.7e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 DIQMTQSPSSLSAS 34  
| | | | | | | | | |  
Db 1 DIQMTQSPSSLSAS 14

Search completed: August 30, 2003, 22:10:22  
Job time : 9.75839 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 22:06:06 ; Search time 22.3825 Seconds  
(without alignments)  
1671.732 Million cell updates/sec

Title: US-08-836-455-2

Perfect score: 145

Sequence: 1 MGAPQILGFLLLFFGTRC.....IKRDAAPTVISIPPSKLG 145

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- SPTREMBL\_23:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriap:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	86.2	127	11 Q925S9	Q925S9 mus musculus
2	28	19.3	239	11 Q8VC55	Q8VC55 mus musculus
3	28	19.3	239	11 Q8K0F8	Q8K0F8 mus musculus
4	26	17.9	233	11 Q91WS9	Q91WS9 mus musculus
5	26	17.9	238	11 Q8VC16	Q8VC16 mus musculus
6	22	15.2	214	11 Q9RIA5	Q9RIA5 mus musculus
7	22	15.2	234	11 Q8R062	Q8R062 mus musculus
8	22	15.2	234	11 Q8VCP0	Q8VCP0 mus musculus
9	22	15.2	238	11 Q9W37	Q9W37 mus musculus
10	19	13.1	112	11 Q8K1F3	Q8K1F3 mus musculus
11	17	11.7	112	11 Q8K1F0	Q8K1F0 mus musculus
12	17	11.7	114	11 Q8K1F1	Q8K1F1 mus musculus
13	17	11.7	134	11 Q8VDD0	Q8VDD0 mus musculus
14	16	11.0	234	11 Q8R028	Q8R028 mus musculus
15	16	11.0	234	11 Q91WF8	Q91WF8 mus musculus
16	16	11.0	235	11 Q91W12	Q91W12 mus musculus

17	15	10.3	298	11 Q9QYF0	Q9QYF0 mus musculus
18	14	9.7	107	4 Q9UL81	Q9UL81 homo sapien
19	14	9.7	107	4 Q96SA9	Q96SA9 homo sapien
20	14	9.7	108	4 Q9UL70	Q9UL70 homo sapien
21	14	9.7	108	4 Q9UL77	Q9UL77 homo sapien
22	14	9.7	108	11 Q8VIJ0	Q8VIJ0 mus musculus
23	14	9.7	116	4 Q96PF6	Q96PF6 homo sapien
24	13	9.0	104	11 Q9JL82	Q9JL82 mus musculus
25	13	9.0	109	11 Q920E6	Q920E6 mus musculus
26	13	9.0	111	11 Q920E9	Q920E9 mus musculus
27	12	8.3	97	11 Q9JL76	Q9JL76 mus musculus
28	12	8.3	101	11 Q9JL78	Q9JL78 mus musculus
29	12	8.3	241	11 Q921A6	Q921A6 mus musculus
30	11	7.6	99	11 Q9JL74	Q9JL74 mus musculus
31	11	7.6	103	11 Q9JL80	Q9JL80 mus musculus
32	11	7.6	107	11 Q9JL84	Q9JL84 mus musculus
33	10	6.9	170	11 Q925S2	Q925S2 mus musculus
34	9	6.2	527	16 Q8DBN7	Q8DBN7 vibrio vuln
35	8	5.5	15	11 Q9QV17	Q9QV17 rattus sp.
36	8	5.5	107	11 Q9ER29	Q9ER29 mus musculus
37	8	5.5	162	6 Q9N1G8	Q9N1G8 salmieri bol
38	8	5.5	220	16 Q8DS60	Q8DS60 streptococc
39	8	5.5	236	16 Q9XA92	Q9XA92 streptomyce
40	8	5.5	239	4 Q8TCD0	Q8TCD0 homo sapien
41	8	5.5	246	8 Q9XPG4	Q9XPG4 leiocassis
42	8	5.5	246	8 Q9XPG5	Q9XPG5 leiocassis
43	8	5.5	261	16 Q8COT1	Q8COT1 staphylococ
44	8	5.5	316	12 Q9OMH4	Q9OMH4 parvo-like
45	8	5.5	345	11 Q8CLH2	Q8CLH2 mus musculus

ALIGNMENTS

RESULT 1

Q925S9 PRELIMINARY; PRT; 127 AA.  
ID Q925S9  
AC Q925S9;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Immunoglobulin light chain (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c;  
RX MEDLINE=99306687; PubMed=10380019;  
RA Tripathi P.K., Qin H., Bhattacharya-Chatterjee M., Ceriani R.L.,  
RA Foon K.A., Chatterjee S.K.;  
RT "Construction and characterization of a chimeric fusion protein  
RT consisting of an anti-idiotypic antibody mimicking a breast cancer-  
RL Hybridoma 18:193-202(1999).  
DR EMBL; AF124721; AAK55120.1;  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IG\_LIKE; 1.  
FT NON\_TER 127  
SQ SEQUENCE 127 AA; 13794 MW; 13F61BEBB981FA5 CRC64;

Query Match 86.2%; Score 125; DB 11; Length 127;

Best Local Similarity 100.0%; Pred. No. 6.4e-122;

Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 APAQILGFLLLFFGTRCDTQMTQSPSSLSASLCQRVSLTCSRASQDGINLHWLQQEPPDG 62

Db 3 APAQILGFLLLFFGTRCDTQMTQSPSSLSASLCQRVSLTCSRASQDGINLHWLQQEPPDG 62

QY 63 TIKRLIYATSLGSGVPRFSGSRGSDYSLTISSEDFVAYICLQYASSPTFFGGT 122  
DB 63 TIKRLIYATSLGSGVPRFSGSRGSDYSLTISSEDFVAYICLQYASSPTFFGGT 122

QY 123 KLEIK 127  
DB 123 KLEIK 127

RESULT 2  
Q8VC55 PRELIMINARY; PRT; 239 AA.  
AC Q8VC55;  
DT 01-MAR-2002 (TReMBLrel. 20, Created)  
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Hypothetical 26.3 kDa protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Colon;  
RA Strausberg R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC021781; AAH21781.1; -.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00406; Ig; 1.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 239 AA; 26303 MW; C16119CACA25C337 CRC64;

Query Match 19.3%; Score 28; DB 11; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.1e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 115 PYTGGGKLEIKRADAAPTIVSIFPPSS 142  
DB 120 PYTGGGKLEIKRADAAPTIVSIFPPSS 147

RESULT 3  
Q8K0F8 PRELIMINARY; PRT; 239 AA.  
AC Q8K0F8;  
DT 01-OCT-2002 (TReMBLrel. 22, Created)  
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Breast tumor;  
RA Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC031498; AAH31498.1; -.  
DR InterPro; IPR002198; ADH\_short.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003597; Ig\_c1.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00409; Ig; 2.  
DR SMART; SM00407; IGc1; 1.  
DR SMART; SM00406; IGv; 1.  
DR PROSITE; PS00061; ADH\_SHORT; 1.

DR PROSITE; PS00290; IG\_MHC; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 239 AA; 26366 MW; D7BE84398AA341FO CRC64;  
Query Match 19.3%; Score 28; DB 11; Length 239;  
Best Local Similarity 100.0%; Pred. No. 1.1e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 PYTGGGKLEIKRADAAPTIVSIFPPSS 142  
DB 120 PYTGGGKLEIKRADAAPTIVSIFPPSS 147

RESULT 4  
Q91WS9 PRELIMINARY; PRT; 233 AA.  
AC Q91WS9;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Hypothetical 25.8 kDa protein (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Colon;  
RA Strausberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC013496; AAH13496.1; -.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00406; Ig; 1.  
DR PROSITE; PS00835; IG\_LIKE; 2.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 233 AA; 25781 MW; B1C184DA149A16EB CRC64;

Query Match 17.9%; Score 26; DB 11; Length 233;  
Best Local Similarity 100.0%; Pred. No. 1.3e-18;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 117 TFGGKLEIKRADAAPTIVSIFPPSS 142  
DB 116 TFGGKLEIKRADAAPTIVSIFPPSS 141

RESULT 5  
Q8VC16 PRELIMINARY; PRT; 238 AA.  
AC Q8VC16;  
DT 01-MAR-2002 (TReMBLrel. 20, Created)  
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Hypothetical 26.2 kDa protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Colon;  
RA Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC019760; AAH19760.1; -.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00406; Ig; 1.



DR PROSITE; PS00290; IG\_MHC; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 238 AA; 26224 MW; 35EC08E3DE5414AD CRC64;

Query Match 17.9%; Score 26; DB 11; Length 238;  
Best Local Similarity 100.0%; Pred. No. 1.3e-18;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 TFGGKLEIKRADAAPTVSIFPPSS 142  
|||||  
DB 121 TFGGKLEIKRADAAPTVSIFPPSS 146

## RESULT 6

Q9RIA5 ID Q9RIA5 PRELIMINARY; PRT; 214 AA.  
AC Q9RIA5  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Kappa light chain of Mab7 (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;  
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal  
RT antibody (Mab 7, its light and heavy chains) and construction of a  
RT single chain antibody (scFv).";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF152371; RAD40242.1; -  
DR HSSP; P01679; 2FBJ.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003006; IG\_MHC.  
DR InterPro; IPR003596; IG\_V.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IG\_LIKE; 2.  
DR PROSITE; PS00290; IG\_MHC; 1.  
FT NON\_TER 1  
FT NON\_TER 214  
SQ SEQUENCE 214 AA; 23922 MW; 52BA205FDE995E2A CRC64;

Query Match 15.2%; Score 22; DB 11; Length 214;  
Best Local Similarity 100.0%; Pred. No. 1.8e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GTKLEIKRADAAPTVSIFPPSS 142  
|||||  
DB 101 GTKLEIKRADAAPTVSIFPPSS 122

## RESULT 7

Q8R062 ID Q8R062 PRELIMINARY; PRT; 234 AA.  
AC Q8R062  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical 25.9 kDa protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RA Strausberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC027418; AAH27418.1; -  
DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003006; IG\_MHC.  
DR InterPro; IPR003596; IG\_V.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IG\_LIKE; 2.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 234 AA; 25857 MW; 4EB08C81426AEAB1 CRC64;

Query Match 15.2%; Score 22; DB 11; Length 234;  
Best Local Similarity 100.0%; Pred. No. 1.9e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GTKLEIKRADAAPTVSIFPPSS 142  
|||||  
DB 121 GTKLEIKRADAAPTVSIFPPSS 142

## RESULT 8

Q8VCP0 ID Q8VCP0 PRELIMINARY; PRT; 234 AA.  
AC Q8VCP0  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical 25.7 kDa protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RA Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC019474; AAH19474.1; -  
DR InterPro; IPR003006; IG\_MHC.  
DR InterPro; IPR003596; IG\_V.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 234 AA; 25702 MW; 102551C58AC2FA9F CRC64;

Query Match 15.2%; Score 22; DB 11; Length 234;  
Best Local Similarity 100.0%; Pred. No. 1.9e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GTKLEIKRADAAPTVSIFPPSS 142  
|||||  
DB 121 GTKLEIKRADAAPTVSIFPPSS 142

## RESULT 9

Q99M37 ID Q99M37 PRELIMINARY; PRT; 238 AA.  
AC Q99M37  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical 26.3 kDa protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC002035; AAH02035.1; -  
DR HSSP; P01679; 2FBJ.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; IG\_MHC.

DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 2.  
 DR SMART: SM00406; IGV; 1.  
 DR PROSITE: PS00835; IG\_LIKE; 2.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 238 AA; FB2B06A0B801330A CRC64;

Query Match 15.2%; Score 22; DB 11; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-14;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GTKLEIKRAADAPTIVSIFPPSS 142.  
 |||||  
 DB 125 GTKLEIKRAADAPTIVSIFPPSS 146

## RESULT 10

Q8K1F3 PRELIMINARY; PRT; 112 AA.  
 AC Q8K1F3;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Anti-VIPase light chain variable region (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/c; TISSUE=Hyperimmunized spleen;  
 RA Zhou Y.-X., Taguchi H., Planque S., Karle S., Nishiyama Y., Paul S.;  
 RT "Innate proteolytic antibodies: Failed D-VIPase response to the D-  
 RT entactinomer of VIP and identification of L-VIPase VL domains.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF516282; AAM64200.1;  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00409; IGV; 1.  
 DR SMART: SM00406; IGV; 1.  
 DR PROSITE: PS00835; IG\_LIKE; 1.  
 FT NON\_TER 1  
 FT NON\_TER 112  
 SQ SEQUENCE 112 AA; 12017 MW; 22DD75B5414CDB18 CRC64;

Query Match 13.1%; Score 19; DB 11; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 PYTFGGGKLEIKRAADAP 133  
 |||||  
 DB 94 PYTFGGGKLEIKRAADAP 112

## RESULT 11

Q8K1F0 PRELIMINARY; PRT; 112 AA.  
 AC Q8K1F0;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Anti-VIPase light chain variable region (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/c; TISSUE=Hyperimmunized spleen;

RA Zhou Y.-X., Taguchi H., Planque S., Karle S., Nishiyama Y., Paul S.;  
 RT "Innate proteolytic antibodies: Failed D-VIPase response to the D-  
 RT entactinomer of VIP and identification of L-VIPase VL domains.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF516285; AAM64203.1;  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00409; IGV; 1.  
 DR SMART: SM00406; IGV; 1.  
 DR PROSITE: PS00835; IG\_LIKE; 1.  
 FT NON\_TER 1  
 FT NON\_TER 112  
 SQ SEQUENCE 112 AA; 11901 MW; F6644663201AA239 CRC64;

Query Match 11.7%; Score 17; DB 11; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 TFGGGTKLEIKRAADAP 133  
 |||||  
 DB 96 TFGGGTKLEIKRAADAP 112

## RESULT 12

Q8K1F1 PRELIMINARY; PRT; 114 AA.  
 AC Q8K1F1;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Anti-VIPase light chain variable region (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/c; TISSUE=Hyperimmunized spleen;  
 RA Zhou Y.-X., Taguchi H., Planque S., Karle S., Nishiyama Y., Paul S.;  
 RT "Innate proteolytic antibodies: Failed D-VIPase response to the D-  
 RT entactinomer of VIP and identification of L-VIPase VL domains.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF516284; AAM64202.1;  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00409; IGV; 1.  
 DR SMART: SM00406; IGV; 1.  
 DR PROSITE: PS00835; IG\_LIKE; 1.  
 FT NON\_TER 1  
 FT NON\_TER 114  
 SQ SEQUENCE 114 AA; 12162 MW; 8BD9833DBF3EEFD1 CRC64;

Query Match 11.7%; Score 17; DB 11; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 TFGGGTKLEIKRAADAP 133  
 |||||  
 DB 98 TFGGGTKLEIKRAADAP 114

## RESULT 13

Q8VDD0 PRELIMINARY; PRT; 134 AA.  
 AC Q8VDD0;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

```
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Anti-MOG 212 variable light chain (Fragment).
GN ANTI-MOG KAPPA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/c;
RA Chernajovsky Y.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/c;
RA Sembl P.;
RL "targeting T cells to the CNS.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ416331; CAC94866.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
FT NON_TER 134
SQ SEQUENCE 134 AA; 14525 MW; CFDF8E2236E2D0CF CRC64;

Query Match 11.78; Score 17; DB 11; Length 134;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 117 TFGGGTKLEIKRADAAP 133
Db 118 TFGGGTKLEIKRADAAP 134

RESULT 14
Q8R028
AC Q8R028; PRELIMINARY; PRT; 234 AA.
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical 25.7 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028540; AAH28540.1; -.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00406; Igv; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 234 AA; 25702 MW; 148377F9C1CD0AEE CRC64;

Query Match 11.0%; Score 16; DB 11; Length 234;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 127 KRADAAPTVSIFPPSS 142
Db 127 KRADAAPTVSIFPPSS 142

RESULT 15
```

```
Q91WF8
AC Q91WF8; PRELIMINARY; PRT; 234 AA.
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical 25.9 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015292; AAH15292.1; -.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR001865; Ribosomal_S2.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00406; Igv; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 234 AA; 25929 MW; B0D0B0E6EB7812D2 CRC64;

Query Match 11.0%; Score 16; DB 11; Length 234;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 127 KRADAAPTVSIFPPSS 142
Db 127 KRADAAPTVSIFPPSS 142

Search completed: August 30, 2003, 22:11:18
Job time : 23.3825 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 30, 2003, 22:10:26 ; Search time 2967.15 Seconds  
(without alignments)  
1999.191 Million cell updates/sec

Title: US-08-836-455-2

Perfect score: 740

Sequence: 1 MGAQAIGLILLFPGR.....IKRADAPTSTFPPSKLG 145

Scoring table:

BLASUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cn2.1/USPTO.spool/US08836455/runat\_29082003.132901.22281/app\_query.fasta\_1.654  
-DB=GenEmbl -QFMT=fastp -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US08836455.ecgn.1.1.3608.etrnat\_29082003.132901.22281 -NCPU=6 -ICPU=3  
-NO\_MMMap -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pin.\*  
35: em.htg\_rod.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em.sv.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	740	100.0	435	6	AR164505	AR164505	Sequence
2	740	100.0	435	6	BD085737	BD085737	Methods o
3	642	86.8	381	10	AF124721	AF124721	Mus muscu
4	633	85.5	407	6	E54981	E54981	Peptide. 1/
5	601	81.2	390	10	MUSIKCC	L41880	Mus musculu
6	599	80.9	381	10	AF045508	AF045508	Mus muscu
7	599	80.9	384	10	AB017434	AB017434	Mus muscu
8	592	80.0	381	10	AF045495	AF045495	Mus muscu
9	581	78.5	381	10	AF045510	AF045510	Mus muscu
10	551	74.5	405	10	AB016620	AB016620	Mus muscu
11	547	73.9	684	10	MUSIGKAC1	J00565	Mus musculu
12	542	73.2	685	10	MMIGK7	V00808	Part of the
13	541	73.1	380	10	MMIGGVJ1	X02177	M.musculus
14	539	72.8	381	10	MMIGGVJ2	X02178	M.musculus
15	533	72.0	354	10	AB089681	AB089681	Mus muscu
16	527	71.2	387	6	AR169918	AR169918	Sequence
17	514	69.5	324	6	AX722008	AX722008	Sequence
18	513	69.3	1019	10	BC027418	BC027418	Mus muscu
19	512	69.2	642	10	AF178454	AF178454	Mus muscu
20	511	69.1	321	10	AF163749	AF163749	Mus muscu
21	511	69.1	383	10	MUSIGKCM	M12191	Mouse Ig ac
22	511	69.1	972	10	AF466770	AF466770	Mus muscu
23	507	68.5	739	6	AR007981	AR007981	Sequence
24	507	68.5	739	6	AR058996	AR058996	Sequence
25	507	68.5	739	6	E54967	E54967	Peptide. 1/
26	506	68.4	324	6	I03643	I03643	Sequence 4
27	506	68.4	324	6	I07835	I07835	Sequence 4
28	506	68.4	348	10	MMVJIG	X54755	Mouse reatr
29	506	68.4	959	10	BC015292	BC015292	Mus muscu
30	506	68.4	959	10	MUSIGRAA3	M59920	Mouse IG ge
31	505	68.2	756	6	AX256284	AX256284	Sequence
32	505	68.2	1497	6	AX256302	AX256302	Sequence
33	503	68.0	432	6	AI7965	AI7965	Variable re
34	503	68.0	438	6	BD015544	BD015544	Human mon
35	503	68.0	438	6	BD094922	BD094922	Human mon
36	503	68.0	432	6	AI7963	AI7963	Variable re
37	502	67.8	711	6	E54975	E54975	Peptide. 1/
38	502	67.8	729	6	E54976	E54976	Peptide. 1/
39	502	67.8	771	6	AX256296	AX256296	Sequence
40	502	67.8	771	6	AX256296	AX256296	Sequence
41	501	67.7	324	10	MUSX	L48667	Mus musculu
42	501	67.7	998	9	S65921	S65921	anti-colore
43	499	67.4	203877	10	AC122260	AC122260	Mus muscu
44	497	67.2	570	10	AF003293	AF003293	Mus muscu
45	496	67.0	330	10	MMIGGIVL	X82890	M.musculus

ALIGNMENTS

RESULT 1

AR164505 LOCUS AR164505 435 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 1 from patent US 6274143.  
ACCESSION AR164505  
VERSION AR164505.1 GI:16237555  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 435)  
AUTHORS Chatterjee, M. and Foon, K.A.  
TITLE Methods of delaying development of HMFg-associated tumors using anti-idiotypic antibody 11D10  
JOURNAL Patent: US 6274143-A 1 14-AUG-2001;  
FEATURES  
source  
1..435  
/organism="unknown"  
BASE COUNT 100 a 111 c 102 g 122 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 7,64e-66 Length: 435  
Score: 740.00 Matches: 145  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-08-836-455-2 (1-145) x AR164505 (1-435)  
QY 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
Db 1 ATGGGGGCCCTCTCAGATTCTTGGGTTCTTGTCTTGTTCAGGTACAGATGT 60  
QY 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
Db 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGATGAGT 120  
QY 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeuGlnGluPro 60  
Db 121 CTCACCTGTGCGGCAAGTCAGGACATGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180  
QY 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
Db 181 GATGGAACTATTAAACGCCCTGATCTACGCCACATCCAGTTAGTCTTGTGTGCCCAAA 240  
QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
Db 241 AGGTTCAAGTGGCAGTAGTCTGGTCCAGATTATCTCTACCATCAGCAGCCTTGAGTCT 300  
QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
Db 301 GAAGATTTTGTAGCTATTACTCTCTACAAATATGCTAGTCTCCGTACACGTTCCGAGGG 360  
QY 121 GlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSerIlePheProPro 140  
Db 361 GGGACCAAGCTGGAATAAAGCGGTGATGCTGCACCAACTGATCTATCCATCTTCCCAACCA 420  
QY 141 SerSerLysLeuGly 145  
Db 421 TCCAGTAAGCTTGGG 435  
RESULT 2  
BD085737  
LOCUS BD085737 435 bp DNA linear PAT 27-AUG-2002  
DEFINITION Methods of delaying development of HMFg-associated tumors using anti-idiotypic antibody 11D10.  
ACCESSION BD085737  
VERSION BD085737.1 GI:22631347  
KEYWORDS JP 2001523269-A/1.  
SOURCE unidentified  
ORGANISM unclassified.

1 (bases 1 to 435)  
Chatterjee, M. and Foon, K.A.  
TITLE Methods of delaying development of HMFg-associated tumors using anti-idiotypic antibody 11D10  
JOURNAL Patent: JP 2001523269-A 1 20-NOV-2001;  
COMMENT THE UNIVERSITY OF KENTUCKY RESEARCH FOUNDATION  
OS Unidentified  
PN JP 2001523269-A/1  
PD 20-NOV-2001  
PF 12-JUN-1998 JP 1999503252  
PR 13-JUN-1997 US 60/049540, 11-JUN-1998 US 09/096244 PI  
PC A61K39/395, A61K39/39//C07K16/42  
CC Strandedness: Single;  
CC Topology: Linear;  
CC Methods of delaying development of HMFg-associated tumors CC  
using  
CC anti-idiotypic antibody 11D10  
FH Key Location/Qualifiers  
FT CDS 1..435  
FT mat peptide 61.  
FEATURES  
source  
1..435  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
BASE COUNT 100 a 111 c 102 g 122 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 7,64e-66 Length: 435  
Score: 740.00 Matches: 145  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-08-836-455-2 (1-145) x BD085737 (1-435)  
QY 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
Db 1 ATGGGGGCCCTCTCAGATTCTTGGGTTCTTGTCTTGTTCAGGTACAGATGT 60  
QY 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
Db 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGACAAAGATGAGT 120  
QY 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeuGlnGluPro 60  
Db 121 CTCACCTGTGCGGCAAGTCAGGACATGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180  
QY 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
Db 181 GATGGAACTATTAAACGCCCTGATCTACGCCACATCCAGTTAGTCTTGTGTGCCCAAA 240  
QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
Db 241 AGGTTCAAGTGGCAGTAGTCTGGTCCAGATTATCTCTACCATCAGCAGCCTTGAGTCT 300  
QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
Db 301 GAAGATTTTGTAGCTATTACTCTCTACAAATATGCTAGTCTCCGTACACGTTCCGAGGG 360  
QY 121 GlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSerIlePheProPro 140  
Db 361 GGGACCAAGCTGGAATAAAGCGGTGATGCTGCACCAACTGATCTATCCATCTTCCCAACCA 420  
QY 141 SerSerLysLeuGly 145  
Db 421 TCCAGTAAGCTTGGG 435  
RESULT 3  
AF124721

```

LOCUS       AF124721               381 bp      mRNA      linear      ROD 22-MAY-2001
DEFINITION  Mus musculus immunoglobulin light chain mRNA, partial cds.
ACCESSION   AF124721
VERSION     AF124721.1   GI:14164546
KEYWORDS    Mus musculus (house mouse)
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE   1 (bases 1 to 381)
AUTHORS     Tripathi,P.K., Qin,H., Bhattacharya-Chatterjee,M., Ceriani,R.L.,
            Foon,K.A. and Chatterjee,S.K.
TITLE       Construction and characterization of a chimeric fusion protein
            consisting of an anti-idiotypic antibody mimicking a breast
            cancer-associated antigen and the cytokine GM-CSF
JOURNAL     Hybridoma 18 (2), 193-202 (1999)
MEDLINE     99306687
PUBMED      10380019
REFERENCE   2 (bases 1 to 381)
AUTHORS     Chatterjee,S.K. and Tripathi,P.K.
TITLE       Direct Submission
JOURNAL     Submitted (29-JAN-1999) Internal Medicine, University of Kentucky,
            800 Rose Street, Lexington, KY 40536, USA
FEATURES    Location/Qualifiers
             source
               1..381
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="BALB/c"
               /db_xref="taxon:10090"
             CDS
               1..>381
               /note="anti-idiotypic antibody 1LD10; mimics a breast
               cancer-associated antigen, human fat globule (HMFG)"
               /codon_start=1
               /product="immunoglobulin light chain"
               /protein_id="AAK55120.1"
               /db_xref="GI:14164547"
               /translation="MRAPQILGFLLLPFGTRCDIOMTQSPSLASLQGRVSLTCR
               ASQDIGNHLWQEPDGTIRLIYATSLSGVPRFSGRSGLSYLTISSESD
               FVAYICQYASSPTFGGTRKLEIK"
BASE COUNT  90 a   93 c   89 g   109 t
ORIGIN
Alignment Scores:
Pred. No.:      5,21e-56      Length:      381
Score:          642.00      Matches:      126
Percent Similarity: 99.21%      Conservative: 0
Best Local Similarity: 99.21%      Mismatches: 1
Query Match:     86.76%      Indels:      0
DB:              10      Gaps:      0

US-08-836-455-2 (1-145) x AF124721 (1-381)

Qy      1  MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuLeuPheProGlyThrArgCys 20
Db      1  ATGAGGGCCCTGCTCAGATCTTGCTTCTGTTCTGTTCCAGGTACCAGATGT 60
Qy      21  AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40
Db      61  GACATCCAGATGACCCAGTCTCCATCTCCTTATCTGCCCTCTCTGGGACAAAGATCAGT 120
Qy      41  LeuThrCysArgAlaSerGlnAspileGlyIleAsnLeuHisTrpLeuGlnGlnGluPro 60
Db      121  CTCACCTTGTGGGCAAGTCAGGACATTCACCATCCAGTTAGTCTTGGTGTCGCCAA 180
Qy      61  AspGlyThrIleLysArgLeuIleTyrAlaThrSerLeuGlySerGlyValProlys 80
Db      181  GATGGAACTATTAACCGCTGATCTACGCCACATCCAGTTAGTCTTGGTGTCGCCAA 240
Qy      81  ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100
Db      241  AGGTTACAGTGCAGTAGTCTGGGTGTCAGATTTATCTCTACCATCAGCAGCCTTGAGTCT 300
Qy      101  GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120

```

```

|||||
301  GAAGATTGTAGCTATTACTGTCTCAATATGTAGTCTTCGTCACACCTGGAGGG 360
Qy      121  GlyThrLysLeuGluIleLys 127
Db      361  GGGACCAAGCTGGAAATAAAA 381

RESULT 4
E54981
LOCUS       E54981               407 bp      DNA      linear      PAT 31-JAN-2002
DEFINITION  Peptide.
ACCESSION   E54981
VERSION     E54981.1   GI:18629719
KEYWORDS    JP 2000236884-A/15.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
            Nishida,T., Okura,T., Tanimoto,T. and Kurimoto,M.
REFERENCE   1 (bases 1 to 407)
AUTHORS     Patent: JP 2000236884-A 15 05-SEP-2000;
            HAYASHIBARA BIOCHEM LAB INC
TITLE       Peptide
JOURNAL
COMMENT     OS Mus musculus (mouse)
            PN JP 2000236884-A/15
            PD 05-SEP-2000
            PF 24-JUN-1999 JP 1999177846
            PR
            PI TAKEHIRO NISHIDA,TAKANORI OKURA,TADAO TANIMOTO, PI MASASHI
            KURIMOTO
            PC C12N15/09,A61K31/00,A61K39/395,A61K48/00,C07K16/24,C12P21/08,
            CC C12N15/00
            CC
            FH Key Location/Qualifiers
            FT CDS (1)..(407)
            FT sig_peptide (1)..(60).
FEATURES    Location/Qualifiers
             source
               1..407
               /organism="Mus musculus"
               /mol_type="genomic DNA"
               /db_xref="taxon:10090"
BASE COUNT  96 a   99 c   96 g   116 t
ORIGIN
Alignment Scores:
Pred. No.:      4,55e-55      Length:      407
Score:          633.00      Matches:      125
Percent Similarity: 94.07%      Conservative: 2
Best Local Similarity: 92.59%      Mismatches: 8
Query Match:     85.54%      Indels:      0
DB:              6      Gaps:      0

US-08-836-455-2 (1-145) x E54981 (1-407)

Qy      1  MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuLeuPheProGlyThrArgCys 20
Db      1  ATGAGGGCCCTGCTCAGATTTTGGCTTCTTGTGTTCTTGTTCCTGTTCCAGGTACCAGATGT 60
Qy      21  AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40
Db      61  GACATCCAGATGACCCAGTCTCCATCTCCTTATCTGCCCTCTCTGGGAGAAAGATCAGT 120
Qy      41  LeuThrCysArgAlaSerGlnAspileGlyIleAsnLeuHisTrpLeuGlnGlnGluPro 60
Db      121  CTCACCTTGTGGGCAAGTCAGGACATTCAGTAAATATATATGCTTCAACAGGAACCA 180
Qy      61  AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProlys 80
Db      181  GATGGAACTTTAAACGCTGATCTACGCCACATCCAGTTAGATTTAGTCTTGGTGTCGCCAA 240
Qy      81  ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100
Db      241  AGGTTACAGTGCAGTAGTCTGGGTGTCAGATTTATCTCTACCATCAGCAGCCTTGAGTCT 300

```

US-08-836-455-2 (1-145) x MUSIKCC (1-390)



```

Alignment Scores:
Pred. No.: 1.15e-51 Length: 381
Score: 599.00 Matches: 118
Percent Similarity: 96.06% Conservatives: 4
Best Local Similarity: 92.91% Mismatches: 5
Query Match: 80.95% Indels: 0
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x AF045508 (1-381)

QY 1 MetGlyAlaProGlnLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20
DB 1 ATAGGGCCCTCCACAGATTTTGGCTTCTTGTCTCTGTTCCAGGTACAGATGT 60
QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40
DB 61 GACATCCAGATGACCCAGTCTCCATCTCTATCTGCCCTCTGGGAGAAAGAGTCAGT 120
QY 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGluPro 60
DB 121 CTCACCTGTGGGCAAGTCAGGACATTTAGTGTAACTGGCTTCAGCAGGAACCA 180
QY 61 AspGlyThrIleLysArgLeuIleTyAlaThrSerSerLeuGlySerGlyValProLys 80
DB 181 GATGGAACTATTAAACGCCGTGATCTAGCCACATCCAGTTAGATTCTGGTGTCACCA 240
QY 81 ArgPheSerGlySerArgSerGlySerAspTyAlaThrSerLeuThrIleSerSerLeuGluSer 100
DB 241 AGGTTTCAGTGGCAGTAGGCTGGGTGATGATTTCTCTCACCATCAGCAGCCCTTGAGTCT 300
QY 101 GluAspPheValAlaTyTyCysLeuGlnTyAlaSerSerProTyThrPheGlyGly 120
DB 301 GAAGATTTGTAGACTATTACTGCTACAAATATGCTAGTTCTCCGTGGAGTTCGGTGA 360
QY 121 GlyThrLysLeuGluIleLys 127
DB 361 GGCACCACGCTGGCAATCAAA 381

RESULT 7
AB017434
LOCUS
DEFINITION
Mus musculus mRNA for anti-IL-18 IgG Light chain, clone 125-2H,
partial cds.
ACCESSION
AB017434.1 GI:6683473
VERSION
KEYWORDS
IgG Light chain.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Nishida.Y.
TITLE
Variable region of anti Human IL-18 IgG Light(kappa) chain
JOURNAL
Published Only in Database (2000)
REFERENCE
2 (bases 1 to 384)
Nishida.Y.
Direct Submission
Submitted (02-SEP-1998) Yoshihiro Nishida, Hayashibara Biochemical
Laboratories, Inc., Fujisaki Institute; 675-1, Fujisaki, Okayama,
Okayama 702-8006, Japan (E-mail:fujihgo.harenet.or.jp,
Tel:+81-86-276-3141, Fax:+81-86-276-6885)
FEATURES
Location/Qualifiers
1..384
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="125-2H"
/cell_type="Hybridoma"
1..>384
/codon_start=1
/product="variable region precursor of anti-IL-18 IgG

Light chain"
/protein_id="BAA89228.1"
/db_xref="GI:6683474"
/translation="MRPAQIFGFLLLPFGTRCDIQMTQSPSSLSASLGERVSLTCR
ASQDYGSKLYWQOEDGCFKRLIYATSSLDSDGVPRFSGSRGSDYSLTISSESD
FVDYICLYQASSPYTFGGGFKLAIR"
1..60
sig_peptide
mat_peptide
61..>384
/product="anti-IL-18 IgG Light chain"
BASE COUNT 91 a 92 c 91 g 110 t
ORIGIN

Alignment Scores:
Pred. No.: 1.16e-51 Length: 384
Score: 599.00 Matches: 118
Percent Similarity: 93.75% Conservatives: 2
Best Local Similarity: 92.19% Mismatches: 8
Query Match: 80.95% Indels: 0
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x AB017434 (1-384)

QY 1 MetGlyAlaProGlnLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20
DB 1 ATAGGGCCCTCCACAGATTTTGGCTTCTTGTCTCTGTTCCAGGTACAGATGT 60
QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40
DB 61 GACATCCAGATGACCCAGTCTCCATCTCTATCTGCCCTCTGGGAGAAAGAGTCAGT 120
QY 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGluPro 60
DB 121 CTCACCTGTGGGCAAGTCAGGACATTTAGTGTAAATTTACTGGCTTCACAGGAACCA 180
QY 61 AspGlyThrIleLysArgLeuIleTyAlaThrSerSerLeuGlySerGlyValProLys 80
DB 181 GATGGAACTTTTAAACGCCGTGATCTAGCCACATCCAGTTAGATTCTGGTGTCACCA 240
QY 81 ArgPheSerGlySerArgSerGlySerAspTyAlaThrSerLeuThrIleSerSerLeuGluSer 100
DB 241 AGGTTTCAGTGGCAGTAGGCTGGGTGATGATTTCTCTCACCATCAGCAGCCCTTGAGTCT 300
QY 101 GluAspPheValAlaTyTyCysLeuGlnTyAlaSerSerProTyThrPheGlyGly 120
DB 301 GAAGATTTGTAGACTATTACTGCTACAAATATGCTAGTTCTCCGTGGAGTTCGGTGA 360
QY 121 GlyThrLysLeuGluIleLysArg 128
DB 361 GGCACCACGCTGGCAATCAAAACGG 384

RESULT 8
AF045495
LOCUS
DEFINITION
Mus musculus dC4 anti-poly(dC) monoclonal antibody kappa light
chain variable region, (IgK) mRNA, partial cds.
ACCESSION
AF045495
VERSION
AF045495.1 GI:2906073
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.
TITLE
Anti-DNA antibodies of normal mice immunized with poly(dC) are
structurally similar to natural autoantibodies
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 381)
O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.
AUTHORS
Direct Submission
TITLE
Submitted (02-FEB-1998) Biochemistry, Tufts University School of
Medicine, 136 Harrison Avenue, Boston, MA 02111, USA
JOURNAL
FEATURES
Location/Qualifiers

```

```

source      1. .381
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="6"
/clone="dc4"
/cell_line="spleen cell hybridoma"
<1. .>381
/gene="Igk"
CDS
1. .>381
/gene="Igk"
/codon_start=1
/product="anti-poly(dc) monoclonal antibody kappa light chain"
/db_xref="GI:2906074"
/translation="MRAPAQIFGFLLLFPCTRCDIQMTQSPSSLSASLGERVSLTCR
ASQDIGSLNWLQOEPDGTIKRLIYATSSLDGVPKRFSGSRGSDYSLTISSESD
FVDYICLYASFPFTFGGKLEIK"
sig_peptide 1. .60
V_region    5.84e-51 Length: 381
           592.00 Matches: 117
Percent Similarity: 95.28% Conservativeness: 4
Best Local Similarity: 92.13% Mismatches: 6
Query Match: 80.00% Indels: 0
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x AF045495 (1-381)

Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuLeuPheProGlyThrArgCys 20
Db 1 ATGAGGCTCTCTGCACAGATTTTGGCTCTTGTGTCTTGTTCAGGTACCATGAT 60

Qy 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40
Db 61 GACATCCAGATGACCATCTCCATCTCTTATCTGCTCTCTGGGAGAAAGATCAGT 120

Qy 41 LeuThrCysArgAlaSerGlnAspileGlyIleAsnLeuHisTrpLeuGlnGlnPro 60
Db 121 CTCACCTTGTGGGCAAGTCAGACATTTGGTAACTAACTAACTGGCTTCACGAGAACCA 180

Qy 61 AspGlyThrIleLysArgLeuIleTyraAlaThrSerSerLeuGlySerGlyValProlys 80
Db 181 GATGGAACATATAAACGCCTGATCTACGCCACATCCAGTTTAGATTTCTGGTGTCCCAAA 240

Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100
Db 241 AGGTTCAGTGGCAGTAGTCTCGGTCAGATTAATCTCTCACCATCAGCAGCCTTGAGTCT 300

Qy 101 GluAspPheValAlaTyrtTyrcysGlnGlnTyraAlaSerProTyrtThrPheGlyGly 120
Db 301 GAACATTTGTAGACTATTACTGCTACATATGCAATATGCAAGTTTTCGGTGACGCTCGGTGGA 360

Qy 121 GlyThrLysLeuGluIleLys 127
Db 361 GGCACCAAGCTGGAAATCAAA 381

RESULT 9
AF045510
LOCUS
DEFINITION Mus musculus 6C9 monoclonal antibody kappa light chain variable region, (Igk) mRNA, partial cds.
ACCESSION AF045510

```

---

```

VERSION      AF045510.1 GI:2906103
KEYWORDS     Mus musculus (house mouse)
SOURCE       Mus musculus
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 381)
AUTHORS      O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.
TITLE        Anti-DNA antibodies of normal mice immunized with poly(dc) are
             structurally similar to natural autoantibodies
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 381)
AUTHORS      O'Connor,K.C., Farrell,T.P., Morikawa,A. and Stollar,B.D.
TITLE        Direct Submission
JOURNAL      Submitted (02-FEB-1998) Biochemistry, Tufts University School of
             Medicine, 136 Harrison Avenue, Boston, MA 02111, USA
FEATURES     Location/Qualifiers
             1. .381
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="C57BL/6"
             /db_xref="taxon:10090"
             /chromosome="6"
             /clone="6C9"
             /cell_line="spleen cell hybridoma"
             <1. .>381
             /gene="Igk"
             1. .>381
             /gene="Igk"
             /codon_start=1
             /product="monoclonal antibody kappa light chain"
             /protein_id="AAC04538.1"
             /db_xref="GI:2906104"
             /translation="MRAPAQIFGFLLLFPCTRCDIQMTQSPSSLSASLGERVSLTCR
             ASQDIGSSLNWLQOEPDGTIKRLIYATSSLDGVPKRFSGSRGSDYSLTISSESD
             FVDYICLYASFPFTFGGKLEIK"
             1. .60
             sig_peptide
             61. .345
             /gene="Igk"
             346. .>381
             /gene="Igk"
             /gene="Igk"
             93 a      89 c      87 g      112 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 7.54e-50 Length: 381
Score: 581.00 Matches: 115
Percent Similarity: 92.91% Conservativeness: 3
Best Local Similarity: 90.55% Mismatches: 9
Query Match: 78.51% Indels: 0
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x AF045510 (1-381)

Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuLeuPheProGlyThrArgCys 20
Db 1 ATGAGGCTCTCTGCACAGATTTTGGCTCTTGTGTCTTGTTCAGGTACCATGAT 60

Qy 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40
Db 61 GACATCCAGATGACCATCTCCATCTCTTATCTGCTCTCTGGGAGAAAGATCAGT 120

Qy 41 LeuThrCysArgAlaSerGlnAspileGlyIleAsnLeuHisTrpLeuGlnGlnPro 60
Db 121 CTCACCTTGTGGGCAAGTCAGACATTTGGTAACTAACTAACTGGCTTCACGAGAACCA 180

Qy 61 AspGlyThrIleLysArgLeuIleTyraAlaThrSerSerLeuGlySerGlyValProlys 80
Db 181 GATGGAACATATAAACGCCTGATCTACGCCACATCCAGTTTAGATTTCTGGTGTCCCAAA 240

Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100
Db 241 AGGTTCAGTGGCAGTAGTCTCGGTCAGATTAATCTCTCACCATCAGCAGCCTTGAGTCT 300

Qy 101 GluAspPheValAlaTyrtTyrcysGlnGlnTyraAlaSerProTyrtThrPheGlyGly 120
Db 301 GAACATTTGTAGACTATTACTGCTACATATGCAATATGCAAGTTTTCGGTGACGCTCGGTGGA 360

Qy 121 GlyThrLysLeuGluIleLys 127
Db 361 GGCACCAAGCTGGAAATCAAA 381

RESULT 9
AF045510
LOCUS
DEFINITION Mus musculus 6C9 monoclonal antibody kappa light chain variable region, (Igk) mRNA, partial cds.
ACCESSION AF045510

```

```

Db      241 AGGTTCAAGTGGCAGTAGGCTGGCTCAGATTATTCTCTCACCATCAGCAGCCTTGAGTCT 300
QY      101 GluAspPheValAlaIaTyTyCysLeuGlnTyAlaSerSerProTyThrPheGlyGly 120
Db      301 GAGATTTTGTAGACTATTACTGTCTCAATAATGCTAGTATTTCCATTACGTTGGGCTCG 360
QY      121 GlyThrLysLeuGluIleLys 127
Db      361 GGCACAAAGTTGGAATAAAA 381

RESULT 10
LOCUS   AB016620                405 bp    mRNA    linear    ROD 27-MAR-2002
DEFINITION Mus musculus mRNA for Immunoglobulin light chain variable region, partial cds.
ACCESSION AB016620
VERSION   AB016620.1 GI:3395670
KEYWORDS Immunoglobulin light chain variable region.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
REFERENCE 1
AUTHORS   Arakawa,F., Yamamoto,T., Kanda,H., Watanabe,T. and Kuroki,M.
TITLE     cDNA sequence analysis of monoclonal antibody FU-MK-1 specific for a transmembrane carcinoma-associated antigen, and construction of a mouse/human chimeric antibody
JOURNAL   Hybridoma 18 (2), 131-138 (1999)
MEDLINE   9305680
PUBMED    10380012
REFERENCE 2 (bases 1 to 405)
AUTHORS   Arakawa,F.
TITLE     Direct Submission
JOURNAL   Submitted (30-JUL-1998) Fumiko Arakawa, School of Medicine, Fukuoka University, First Department of Biochemistry; 7-45-1 Nanakuma, Jonan-ku, Fukuoka, Fukuoka 814-80, Japan
          (E-mail:farakawa@emsat.fukuoka-u.ac.jp, Tel:092-801-1011(ex.3246), Fax:092-801-3600)
FEATURES             Location/Qualifiers
     source           1..405
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell_line="FU-MK-1 hybridoma"
                     /note="the hybridoma producing anti-GA733-2 monoclonal antibody"
     exon             1..392
     primer_bind      1..43
                     /note="PCR primer region"
     cds              11..>392
                     /codon_start=1
                     /product="Immunoglobulin light chain variable region"
                     /protein_id="PAA32080.1"
                     /db_xref="GI:3395671"
                     /translat="IMGT/IGM:AB016620"
                     /translation="MRAPAIQLGLLIMFFGIRCDIKMTQSPSSLSASLGERSVSLR ASQEISGLSWLQKQPDGVTKRLIYAASLTHSGVPRKFRSGSGSDYSLTISLESDD
                     FADYICLOYASDPWTFGGTGLEIK"
     sig_peptide      11..70
     intron            393..405
     primer_bind      393..405
                     /note="PCR primer region"
     BASE COUNT      98 a 100 c 97 g 110 t
     ORIGIN
Alignment Scores:
Pred. No.:          8.65e-47          Length:          405
Score:              551.00           Matches:         108
Percent Similarity: 90.08%           Conservative:    10
Best Local Similarity: 82.44%         Mismatches:      13
Query Match:        74.46%           Indels:          0
DB:                  10              Gaps:            0

```

```

US-08-836-455-2 (1-145) x AB016620 (1-405)
QY      1 MetClyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuLeuPheProGlyThrArqCys 20
Db      11 ATGAGGCCCTCTGCTCAGATTCTGGCTTCTTGCTCTGGTTTCCAGGTATAGATGT 70
QY      21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40
Db      71 GACATCAAGATGACCCAGTCGCCATCTCTTATCTGCTCTCTGGGAGAAAGAGTCACT 130
QY      41 LeuThrCysArgAlaSerGlnAspIleGlyIleasnLeuHisTrpLeuGlnGlnGluPro 60
Db      131 CTCACCTGTCTGGCAAGTCAGGAATTAGTGGTTACTTAAGCTGGCTTCAGCAAAACCA 190
QY      61 AspGlyThrIleLysArgLeuIleTyAlaThrSerSerLeuGlySerGlyValProLys 80
Db      191 GATGGAACCTGTTAAACGCCCTGATCTAGCCGCATCCACTTACATCTTGGTGTCCCAAAA 250
QY      81 ArgPheSerGlySerArgSerGlySerAspTySerLeuThrIleSerSerLeuGluSer 100
Db      251 AGGTTCAGTGGCAGTAGGCTGGTGCAGACTATTCTCTCACCATCAGCAGCCTTGAGTCT 310
QY      101 GluAspPheValAlaIaTyTyCysLeuGlnTyAlaSerSerProTyThrPheGlyGly 120
Db      311 GAGCATTTTGCAGACTATTACTGTCTACAGTATGCTAGTATGATCGGTGGAGTTCGGTGA 370
QY      121 GlyThrLysLeuGluIleLysArgAlaAspAla 131
Db      371 GGCACCAAGCTGGNAATCAACGTAAGAATTCG 403

RESULT 11
LOCUS   MUSIGKAC1                684 bp    DNA    linear    ROD 19-JUN-2002
DEFINITION Mus musculus immunoglobulin kappa chain variable region (Igk) gene, exons 1 and 2.
ACCESSION J00565
VERSION   J00565.1 GI:196531
KEYWORDS
SEGMENT    1 of 3
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE 1
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE     Seidman,J.G., Max,E.E. and Leder,P.
JOURNAL   A kappa-immunoglobulin gene is formed by site-specific recombination without further somatic mutation
MEDLINE   Nature 280 (5721), 370-375 (1979)
PUBMED    79221900
REFERENCE 2 (bases 1 to 684)
AUTHORS   Gough,N.M., Cory,S. and Adams,J.M.
TITLE     Identical 3' non-coding sequences in five mouse Ig kappa chain mRNAs favour a unique C kappa gene
JOURNAL   Nature 281 (5730), 394-396 (1979)
MEDLINE   80011674
PUBMED    113684
REFERENCE 3 (sites)
AUTHORS   Queen,C. and Baltimore,D.
TITLE     Immunoglobulin gene transcription is activated by downstream sequence elements
JOURNAL   Cell 33 (3), 741-748 (1983)
MEDLINE   83259260
PUBMED    6409419
COMMENT    [2] sites; comment.
           this sequence is a productively rearranged kappa ig from myeloma mopc41. the conflicts noted in the sites table refer to the published differences between the germline and active genes. they are probably typographical errors since [1] says that no somatic mutation has occurred after recombination. j.g. seidman (personal communication) said that the published germline sequence is correct <musigkvc>. [2] finds that deletion of 1.3kb 5' to the constant region exon results in a lowered rate of transcription and an

```

altering of the site of transcription initiation. the deletion moves transcription initiation from approximately 30 bp 5' to the initiation codon to about 20 bp 5' for part of the vk41 constant gene and part of the 3' flank see <musigkac2> and <musigkac3> respectively. for the germline joining j1 region see <musigkjc2>. for other rearranged kappa genes see loci beginning <musigka>, and for germline kappa variable genes see loci beginning <musigkv>. in the sites table cdr-complementarity determining region and fr-framework region.

FEATURES		Location/Qualifiers	
source	1..684	/organism="Mus musculus"	
		/mol_type="genomic DNA"	
sig_peptide	119..173	/db_xref="taxon:10090"	
		/cell_line="MOPC41"	
exon	174..301	/tissue_type="myeloma"	
		/rearranged	
intron	302..634	join(119..173,302..312)	
		/gene="Igk"	
exon	596..597	<119..173	
		/number=1	
misc_recomb	164 a	/number=2	
		/gene="Igk"	
BASE COUNT	164 a	156 c	147 g 217 t
ORIGIN			
Alignment Scores:			
Pred. No.:	3.94e-46	Length:	684
Score:	547.00	Matches:	118
Percent Similarity:	71.35%	Conservative:	4
Best Local Similarity:	69.01%	Mismatches:	6
Query Match:	73.92%	Indels:	44
DB:	10	Gaps:	1
US-08-836-455-2 (1-145) x MUSIGKAC1 (1-684)			
QY	1	MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPhePro	----- 16
Db	125	ATGAGGCTCTGCACAGATTTTGGCTCTTGTGCTTGTGTTCA-AGGTAAATGAA	183
QY	16	-----	16
Db	184	ACTTAAATTTGGGAATTTCCACGTTCACACTGTTCCAACTGTGTTAGTGTGACTGGCATTTGGG	243
QY	17	-----	17
Db	244	GGATGCTCTCTTTATCATGCTTATCATGTTGATATTATGCTCCACTCTAGGT	303
QY	18	ThrArgCysAspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGln	37
Db	304	ACCAGATGTGACATCCAGATGACCCAGTCTCCATCCCTTATCTGCTCTCTGGGAGAA	363
QY	38	ArgValSerLeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGln	57
Db	364	AGATGCTCTCACTTGTGCGGCAAGTCAGGACATTTGGTAGTACCTTAACTGGCTTCAG	423
QY	58	GlnGluProAspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGly	77
Db	424	CAGGAACAGATGGAACATTATTAACGGCTGATCTACGCCACATCCAGTTAGATTTCTGTT	483
QY	78	ValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSer	97
Db	484	GTCCCCAAAGTTTCAGTGGCAGTAGGTCTGGGGTCAGATTATCTCTCACCATCAGCAGC	543
QY	98	LeuGluSerGluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThr	117
US-08-836-455-2 (1-145) x MMIGK7 (1-685)			
QY	1	MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPhePro	----- 16
Db	126	ATGAGGCTCTGCACAGATTTTGGCTCTTGTGCTTGTGTTCA-AGGTAAATGAA	184
QY	16	-----	16
Db	185	ACTTAAATTTGGGAATTTCCACGTTCACACTGTTCCAACTGTGTTAGTGTGACTGGCATTTGGG	244
QY	17	-----	17
Db	245	GGATGCTCTCTTTATCATGCTTATCATGTTGATATTATGCTCCACTCTAGGT	304
QY	18	ThrArgCysAspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGln	37
Db	305	ACCAGATGTGACATCCAGATGACCCAGTCTCCATCCCTTATCTGCTCTCTGGGAGAA	364
QY	38	ArgValSerLeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGln	57
Db	365	AGATGCTCTCACTTGTGCGGCAAGTCAGGACATTTGGTAGTACCTTAACTGGCTTCAG	424

FEATURES		Location/Qualifiers	
source	1..685	/organism="Mus musculus"	
		/mol_type="genomic DNA"	
sig_peptide	120..173	/db_xref="taxon:10090"	
		join(120..173,302..>637)	
exon	302..637	/codon_start=1	
		/product="immunoglobulin kappa"	
intron	638..1190	/protein_id="CAA24190.1"	
		/db_xref="GI:577074"	
exon	1191..1730	/translation="MDMRAPIQIFGLLLFQGTQDIOIOWTQSPSSLSASLGERVSLT	
		CRPSQDSSGLNMLQEPDGTIKRLIYATSSLDGVPKRFSGRSQSDYSLTISLES	
misc_recomb	164 a	EDFDYYCLOYASSPWTFGGTTKLEIKR"	
		sig_peptide	
BASE COUNT	164 a	158 c	146 g 217 t
ORIGIN			
Alignment Scores:			
Pred. No.:	1.26e-45	Length:	685
Score:	542.00	Matches:	117
Percent Similarity:	70.76%	Conservative:	4
Best Local Similarity:	68.42%	Mismatches:	7
Query Match:	73.24%	Indels:	44
DB:	10	Gaps:	1
US-08-836-455-2 (1-145) x MMIGK7 (1-685)			
QY	1	MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPhePro	----- 16
Db	126	ATGAGGCTCTGCACAGATTTTGGCTCTTGTGCTTGTGTTCA-AGGTAAATGAA	184
QY	16	-----	16
Db	185	ACTTAAATTTGGGAATTTCCACGTTCACACTGTTCCAACTGTGTTAGTGTGACTGGCATTTGGG	244
QY	17	-----	17
Db	245	GGATGCTCTCTTTATCATGCTTATCATGTTGATATTATGCTCCACTCTAGGT	304
QY	18	ThrArgCysAspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGln	37
Db	305	ACCAGATGTGACATCCAGATGACCCAGTCTCCATCCCTTATCTGCTCTCTGGGAGAA	364
QY	38	ArgValSerLeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGln	57
Db	365	AGATGCTCTCACTTGTGCGGCAAGTCAGGACATTTGGTAGTACCTTAACTGGCTTCAG	424

544 CTTGAGTCTGAAGATTTTGTAGACTATTACTGTCTACAATATGCTAGTCTCTCCGTGGACG 603

118 PheGlyGlyGlyThrLysLeuGluIleLysArg 128

604 TTGCGTGGAGCACCAGCTGGAAATCAACGT 636

RESULT 12

MMIGK7

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

FEATURES

source

CDS

sig\_peptide

exon

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-08-836-455-2 (1-145) x MMIGK7 (1-685)

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

```

QY 58 GlnGluProAspGlyThrLeuLysArgLeuLeuIleTyrAlaThrSerSerLeuGlySerGly 77
    |||||
Db 425 CAGAACACAGATGGAACATAAAGCCCTGATCTACGCCACATCCAGTTTAGATTCTGGT 484
QY 78 ValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSer 97
    |||||
Db 485 GTCCCCAAGGTTTCAGTGGCAGTAGGTCTGGGTCTGAGATTATTTCTCTCACCATCAGCAG 544
QY 98 LeuGluSerGluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThr 117
    |||||
Db 545 CTGTGATCTGAGATTTTGTAGACATATTTACTGTCTACAAATATGCTAGTTCTCGGTGGACG 604
QY 118 PheGlyGlyGlyThrLysLeuGluLeuLysArg 128
    |||||
Db 605 TTCGGTGGAGGCCAACCAAGCTGGGAATCAAAACGT 637

RESULT 13
MMIGGVJ1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
FEATURES
    source
        1. .380
            /organism="Mus musculus"
            /mol_type="mRNA"
            /db_xref="taxon:10090"
            <1..>380
            /codon_start=3
            /product="immunoglobulin G kappa light chain"
            /protein_id="CAA26118.1"
            /db_xref="GI:663215"
            /translation="FGFLLWFPPTGTRCDIQMTQSPSSLSASLGERVSLTCRASQETSG
            YLSWLQOKPDGTIKRLIYAASTLDGVPKRFSGRRSGSDYSLTISSEDEADYVCL
            QYLSVPLTFGAGTKLELRADAAP"
V_region
    misc_feature
        42..325
            /product="variable region of IgG kappa light chain"
            111..143
            /notes="hypervariable region CDR1"
    misc_feature
        189..212
            /notes="hypervariable region CDR2"
    misc_feature
        306..332
            /notes="hypervariable region CDR3"
    J_segment
        326..364
            /product="joining region of IgG kappa light chain"
    C_region
        365..>380
            /product="constant region of IgG kappa light chain"
BASE COUNT 88 a 93 c 89 g 110 t
ORIGIN
Alignment Scores:
Pred. No.: 8 24e-46 Length: 380
Score: 541.00 Matches: 107
Percent Similarity: 89.60% Conservative: 5
Best local Similarity: 85.60% Mismatches: 13
Query Match: 73.11% Indels: 0
DB: 10 Gaps: 0
US-08-836-455-2 (1-145) x MMIGGVJ1 (1-380)

```

```

QY 9 GlyPheLeuLeuLeuPheProGlyThrArgCysAspIleGlnMetThrGlnSerPro 28
    |||||
Db 6 GGCTTCTGTGTTGCTCTGTTCCAGGTACCAGATGTGACATCCAGATGACCCAGTCTCCA 65
QY 29 SerSerLeuSerAlaSerLeuGlyGlnArgValSerLeuThrCysArgAlaSerGlnAsp 48
    |||||
Db 66 TCCTCTTATCTCCCTCTCTGGGAGAAAGAGTCAGTCTCACTTGTTCGGGCAAGTCAGGAA 125
QY 49 IleGlyIleAsnLeuHisTrpLeuGlnGlnProAspGlyThrIleLysArgLeuIle 68
    |||||
Db 126 ATTAGTGGTTACTTAGCTGGCTTCAGCAGAAACAGATGGAAGTATTAACCCCTGATC 185
QY 69 TyrAlaThrSerSerLeuGlySerGlyValProLysArgPheSerGlySerArgSerGly 88
    |||||
Db 186 TAGCCGCATCCACTTATAGATTCTGGTGTCCCAAAAGGTTTCAGTGGCAGTAGTCTGGG 245
QY 89 SerAspTyrSerLeuThrIleSerSerLeuGluSerGluAspPheValAlaTyrTyrCys 108
    |||||
Db 246 TCAGATATATCTCTCACCATCAGCAGCCTTGAGTCTGAAGATTTTCAGACTATTACTGT 305
QY 109 LeuGlnTyrAlaSerSerProTyrThrPheGlyGlyThrLysLeuGluLeuLysArg 128
    |||||
Db 306 CTACAATATCTTAGTTATCGCTCAGTTCGGTCTGGGACCAAGCTGGAGCTGAACACGG 365
QY 129 AlaAspAlaAlaPro 133
    |||||
Db 366 GCTGATGCTGCACCA 380

RESULT 14
MMIGGVJ2
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
FEATURES
    source
        1. .381
            /organism="Mus musculus"
            /mol_type="mRNA"
            /db_xref="taxon:10090"
            <1..>381
            /codon_start=1
            /product="immunoglobulin G kappa light chain"
            /protein_id="CAA26119.1"
            /db_xref="GI:663216"
            /translation="VFGFLLWFPPTGTRCDIQMTQSPSSLSASLGERVSLTCRASQETIS
            GYLSWLQOKPDGTIKRLIYAASTLDGVPKRFSGRRSGSDYSLTISSEDEADYVYC
            LOYLSVPLTFGAGTKLELRADAAP"
V_region
    misc_feature
        43..326
            /product="variable region of IgG kappa light chain"
            112..144
            /notes="hypervariable region CDR1"
    misc_feature
        190..213
            /notes="hypervariable region CDR2"
    misc_feature
        307..333
            /notes="hypervariable region CDR3"
    J_segment
        327..365
            /product="joining region of IgG kappa light chain"
    C_region
        366..>381
            /product="constant region of IgG kappa light chain"
BASE COUNT 90 a 92 c 89 g 110 t
ORIGIN

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 30, 2003, 22:10:01 ; Search time 217.987 Seconds  
(without alignments)  
1795.608 Million cell updates/sec

Title: US-08-836-455-2

Perfect score: 740

Sequence: 1 MGAPQAIGLFLLLFPFTRC.....IKRADAAPTVSIFPPSSKLG 145

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh  
-Q/crn2\_1/USPTO\_spool/US08836455/runat\_29082003\_132901\_22271/app\_query.fasta\_1.654  
-DB=N\_Geneseq\_19Jun03 -QFMT=fastp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=ptc -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US08836455.ecgn\_1.1.401@runat\_29082003\_132901\_22271 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBIOCK=100 -LONGLOG  
-DSV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_19Jun03: \*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: \*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: \*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: \*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: \*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: \*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: \*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: \*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: \*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: \*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: \*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: \*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: \*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: \*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: \*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: \*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: \*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: \*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: \*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: \*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: \*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: \*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: \*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: \*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: \*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	740	100.0	435	18	AAT85149	Murine monoclonal
2	740	100.0	435	20	AAV83772	Antibody 11D10 11g
3	740	100.0	435	25	AA51273	Mouse 11D10 antibo
4	662	89.5	450	4	AA30165	Sequence encoding
5	633	85.5	407	21	AA49548	Mouse light chain
6	527	71.2	387	18	AAT77851	Murine anti-human
7	527	71.2	387	24	AA32138	Murine 44H104 mab
8	507	68.5	739	14	AA046084	Sequence encoding
9	507	68.5	739	17	AAT36880	520C9 anti-c-erbB-
10	507	68.5	739	19	AAV21798	520C9 anti-c-erbB-
11	507	68.5	739	20	AAV63399	520C9 sfv DNA sequ
12	506	68.4	324	21	AA49534	Mouse anti-IL-18 a
13	505	68.2	756	24	AA597136	P4-3 single chain
14	505	68.2	1497	24	AA597145	3B10xP4-3 bispecif
15	503	68.0	432	12	AAQ15114	IL-2 chimeric anti
16	503	68.0	438	22	AAH41157	Human coding sequ
17	502	67.8	711	21	AA49542	pESCFv#125-2H reco
18	502	67.8	729	21	AA49543	pESCFv#125-2H.Ht r
19	502	67.8	771	24	AA597142	P5-10 single chain
20	497	67.2	432	12	AAQ15113	IL-2 chimeric anti
21	495	66.9	1605	14	AA046086	Sequence encoding
22	492	66.5	321	21	AAA38908	520C9 hybridoma VL
23	492	66.5	439	18	AAT73443	Human immunoglobul
24	492	66.5	439	20	AAZ21995	Partial nucleotide
25	486	65.7	3819	18	AAT78825	Kappa light chain
26	486	65.7	3819	19	AAV39266	Plasmid pLC665 nuc
27	486	65.7	3819	20	AAZ22820	Nucleotide sequenc
28	484.5	65.5	737	24	AAZ31829	Human pancreatic t
29	483	65.3	449	15	AAQ78732	Murine anti-human
30	482	65.1	817	21	AAA27389	Human IGFAM-9 immu
31	481	65.0	714	21	AAA46899	DNA encoding the k
32	481	65.0	19035	19	AAV61794	Traget plasmid Man
33	479	64.7	456	22	AAQ66996	Filamentous phage
34	479	64.7	867	22	AAQ67002	Filamentous phage
35	479	64.7	1066	14	AAQ49943	Human anti-HBs lig
36	473	63.9	974	24	AA599473	Anti-human AILIM m
37	471	63.6	447	17	AAT34541	Monoclonal anti-id
38	471	63.6	447	18	AAT31540	3H1 light chain va
39	471	63.6	447	17	AAT99434	Anti-idiotypic anti
40	471	63.6	447	21	AAZ35842	Murine MAB against
41	471	63.6	447	22	AAQ09316	Mouse anti-idiotyp
42	471	63.6	447	24	ABL60799	Antibody 3H1 light
43	471	63.6	447	24	AAK98279	Mouse 3H1 antibody
44	468	63.2	321	16	AAQ97504	Light chain variab
45	468	63.2	321	16	AAQ97507	Light chain variab

ALIGNMENTS

RESULT 1

AAT85149

ID AAT85149 standard; cDNA; 435 BP.

AC AAT85149;

DT 25-MAR-2003 (updated)

DT 04-JAN-1998 (first entry)

DE Murine monoclonal anti-idiotypic antibody 11D10 VL cDNA.

KW Monoclonal antibody 11D10; anti-idiotypic antibody; mucin;

KW human milk fat globule; HMFG; tumour; breast cancer; vaccine; ss.

OS Mus musculus.

XX Key

XX Location/Qualifiers

FH





## Alignment Scores:

Pred. No.: 1.81e-61 Length: 435  
 Score: 740.00 Matches: 145  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

US-08-836-455-2 (1-145) x AAV83772 (1-435)

QY 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
 DB 1 ATGGGGCCCCCTCTCAGATCTTGGGTTCTTGTCTTGTTCAGGTACCATGATGT 60  
 QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
 DB 61 GACATCCAGATGACCCAGTCTCATCTCTTATCTGCCTCTCTGGACAAAGAGTCAGT 120  
 QY 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGluPro 60  
 DB 121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180  
 QY 61 AspGlyThrIleLysArgLeuIleTyAlaThrSerSerLeuGlySerGlyValProLys 80  
 DB 181 GATGGAACTATTAAACGCCCTGATACGCCACATCCAGTTAGGTTCTGGTGTCCCAAA 240  
 QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGlySer 100  
 DB 241 AGGTTTCAGTGGCAGTAGGTCGGTCTCAGATTTATCTCTCACCATCAGCAGCCTTGAGTCT 300  
 QY 101 GluAspPheValAlaIleTyThrCysLeuGlnTyAlaSerSerProTyThrPheGlyGly 120  
 DB 301 GAAGATTTGTAGCCCTATTACTCTCTACAAATATGCTAGTTCTCCGTACAGTTCCGGAGGG 360  
 QY 121 GlyThrLysLeuGluIleLysArgAlaAspAlaProThrValSerIlePheProPro 140  
 DB 361 GGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACTGTATCCATCTTCCCAACCA 420

## RESULT 3

AAL51273  
 ID AAL51273 standard; cDNA; 435 BP.

XX AC AAL51273;

XX 20-MAR-2003 (first entry)

XX Mouse 11D10 antibody light chain variable region coding sequence.

DE Mouse; murine; gene; ss; vaccine; tumour; human milk fat globules; HMFG;  
 KW carcinoembryonic antigen; CEA; 11D10; 3H1; HMFG-associated tumour;  
 KW CEA-associated tumour; anti-idiotype antibody.

XX OS Mus musculus.

XX Key Location/Qualifiers

XX CDS 1..435

XX /\*tag= a

XX /partial

XX /product= "Mouse 11D10 anti-idiotype antibody light chain

XX variable region"

XX /note= "No stop codon is given"

XX sig\_peptide 1..60

XX /\*tag= b

XX mat\_peptide 61..435

XX /\*tag= c

XX WO200292012-A2.

XX 21-NOV-2002.

XX PN

XX XX

XX PD

XX

XX 17-MAY-2002; 2002WO-US15840.

XX PF

XX 17-MAY-2001; 2001US-0861294.

XX PR

XX (KENT ) UNIV KENTUCKY RES FOUND.

XX PA

XX Chatterjee M, Foon KA;

XX PI

XX WPI; 2003-129216/12.

XX DR

XX P-PSDB; AAO16292.

XX XX

Use of anti-idiotypic antibodies for human milk fat globules (HMFG)- or  
 carcinoembryonic antigen (CEA)-associated tumor for delaying the  
 development of, or treating a HMFG- or CEA-associated tumor (e.g.  
 breast tumor) in humans

XX Disclosure; Fig 1; 98pp; English.

XX CC

The invention comprises a method for delaying the development of, or  
 treating a tumour that is associated with human milk fat globules (HMFG)  
 or carcinoembryonic antigen (CEA). The method of the invention involves  
 administering an anti-idiotypic antibody for HMFG (e.g. 11D10), and an  
 anti-idiotypic antibody for CEA (e.g. 3H1). The method is useful for  
 delaying the development, of or treating HMFG/CEA-associated tumours. The  
 present cDNA sequence encodes the light chain variable region of the  
 mouse 11D10 anti-idiotype antibody.

XX XX

XX Sequence 435 BP; 100 A; 111 C; 102 G; 122 T; 0 other;

## Alignment Scores:

Pred. No.: 1.81e-61 Length: 435  
 Score: 740.00 Matches: 145  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 25 Gaps: 0

US-08-836-455-2 (1-145) x AAL51273 (1-435)

QY 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
 DB 1 ATGGGGCCCCCTCTCAGATTTCTTGGTCTTGTGTCTTGTTCAGGTACCATGATGT 60

QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40

DB 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCCTCTCTGGACAAAGAGTCAGT 120

QY 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGluPro 60

DB 121 CTCACCTTGTGGGCAAGTCAGGACATTTGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180

QY 61 AspGlyThrIleLysArgLeuIleTyAlaThrSerSerLeuGlySerGlyValProLys 80

DB 181 GATGGAACTATTAAACGCCCTGATCTACGCCACATCCAGTTAGGTTCTGGTGTCCCAAA 240

QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGlySer 100

DB 241 AGGTTTCAGTGGCAGTAGGTCGGTCTCAGATTTATCTCTCACCATCAGCAGCCTTGAGTCT 300

QY 101 GluAspPheValAlaIleTyThrCysLeuGlnTyAlaSerSerProTyThrPheGlyGly 120

DB 301 GAAGATTTGTAGCCCTATTACTGTCTACAAATATGCTAGTTCTCCGTACAGTTCCGGAGGG 360

QY 121 GlyThrLysLeuGluIleLysArgAlaAspAlaProThrValSerIlePheProPro 140

DB 361 GGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACTGTATCCATCTTCCCAACCA 420

QY 141 SerSerLysLeuGly 145

DB 421 TCCAGTAAGCTTGGG 435

RESULT 4

AAN30165  
ID AAN30165 standard; DNA; 450 BP.  
XX  
AC AAN30165;  
XX  
DT 25-MAY-1992 (first entry)  
XX  
DE Sequence encoding the leader, variable region and first 16 AAs of  
DE the constant region of the kappa-chain (light chain) of MOPC41.  
XX  
KW Diagnosis; therapy; immunoglobulin; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 13..78  
FT /\*tag= a  
FT /product= leader  
FT 79..402  
FT CDS /\*tag= b  
FT /product= variable region  
FT 403..450  
FT /\*tag= c  
FT /product= constant region  
XX  
EP88994-A.  
XX  
PD 21-SEP-1983.  
XX  
PF 10-MAR-1983; 83EP-0001655.  
XX  
PR 15-MAR-1982; 82US-0358414.  
PR 05-DEC-1983; 83US-0558551.  
XX  
PA (SCHE ) SCHERING CORP.  
PA (DNAX ) DNAX RES INST.  
XX  
PI Moore KW, Zaffaroni A;  
XX  
DR WPI; 1983-772290/39.  
DR P-PSDB; AAP30251.  
XX  
PT Transformed expression vectors or plasmid(s) - with double  
PT stranded DNA sequence coding only for desired part of polypeptide  
PT chain  
XX  
PS Example; Page 40-41; 68pp; English.  
XX  
CC The pref. vector or plasmid of the invention has a double-stranded  
CC DNA seq. coding for a variable region of a light or heavy chain of  
CC IgG, or for a variable region of a light or heavy chain of an  
CC immunoglobulin specific for an enzyme or surface protein. The  
CC sequence esp. codes for a variable region of a light chain having 95-  
CC 115 AAs or for a variable region of a heavy chain having 110-125 AAs  
CC esp. including the D region of the heavy chain.  
XX  
SQ Sequence 450 BP; 108 A; 114 C; 104 G; 124 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 4,61e-54 Length: 450  
Score: 662.00 Matches: 131  
Percent Similarity: 95.10% Conservative: 5  
Best Local Similarity: 91.61% Mismatches: 7  
Query Match: 89.46% Indels: 0  
DB: 4 Gaps: 0  
  
US-08-836-455-2 (1-145) x AAN30165 (1-450)  
  
Qy 1 MetGlylaProAlaGlnLeuLeuGlyPheLeuLeuLeuLeuPheProGlyThrArgCys 20  
Db 19 ATGAGGGCTCTGTCACAGATTTTGGCTCTTGTGCTCTTGTTCAGGTACCATGATG 78  
Qy 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40

Db 79 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGGAGAAAGATCAGT 138  
Qy 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGlnGluPro 60  
Db 139 CTCACCTTGTGCGCCAAGTCAGGACATTGGTAGTAGCTTAACTGCTTACGACGAGACCA 198  
Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
Db 199 GATGGAACATTAACGCCCTGATCTACGCCACATCCAGTTTAGATTCTGGTGTCCCCAAA 258  
Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
Db 259 AGGTTTCAGTGGCAGTAGTCTGGGTGCAGATTATCTCTCACCATCAGCAGCCTTGAGTCT 318  
Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
Db 319 GAAGATTTTGTAGACTATTACTGTCTACAATATGCTAGTCTCCGTGGACCTCGGTGA 378  
Qy 121 GlyThrLysLeuGluIleLysArgAlaAspAlaProThrValSerIlePheProPro 140  
Db 379 GGCACCAAGCTGGAAATCAACGCTGCTGATCTGCACCAACTGTATCCATCTTCCACCA 438  
Qy 141 SerSerLys 143  
Db 439 TCCAGTGAG 447  
  
RESULT 5  
AAZ49548  
ID AAZ49548 standard; cDNA; 407 BP.  
XX  
AC AAZ49548;  
XX  
DT 04-APR-2000 (first entry)  
XX  
DE Mouse light chain variable region DNA.  
XX  
KW Mouse light chain variable region; VL; variable region light chain;  
KW interleukin-18; immunopathies; inflammatory disorder; autoimmune disease;  
KW mouse; anti-allergic; anti-inflammatory; immunosuppressive;  
KW hematopoietic; leukocytopenic; antialgic; antipyretic; ds.  
XX  
OS Mus musculus.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT mat\_peptide 1..407  
FT /\*tag= a  
FT /product= "Mouse light chain variable region"  
FT sig\_peptide 1..60  
FT /\*tag= b  
XX  
PN EP974600-A2.  
XX  
PD 26-JAN-2000.  
XX  
PF 24-JUN-1999; 99EP-0304977.  
XX  
PR 24-JUN-1998; 98JP-0177580.  
PR 12-OCT-1998; 98JP-0289044.  
PR 22-DEC-1998; 98JP-0365023.  
XX  
PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
XX  
PI Nishida Y, Okura T, Tanimoto T, Kurimoto M;  
XX  
DR WPI; 2000-118341/11.  
DR P-PSDB; AAY44599.  
XX  
PT New artificially produced peptide for neutralizing biological activity  
PT of interleukin-18, useful for treating and preventing immunopathies,  
PT inflammatory disorders and autoimmune diseases -  
XX

PS Example 1; Page 28-29; 32pp; English.

XX The present sequence encodes mouse light chain variable region. This  
 CC recombinant DNA is derived from PCR A which amplifies antibody light  
 CC chain variable region (VL). The transformant produced using the VL gene  
 CC was used transform competent E. coli cells. The peptide produced by  
 CC transformants neutralises interleukin-18. This is useful for treating and  
 CC preventing immunopathies, inflammatory disorders and autoimmune diseases  
 CC which are caused by excessive immunoreaction. The peptide has  
 CC anti-allergic, anti-inflammatory, immunosuppressive, hematopoietic,  
 CC leukocytopenic, antialgic, antipyretic and hepatic-function improving  
 CC activities.

XX SQ Sequence 407 BP; 96 A; 99 C; 96 G; 116 T; 0 other;

Alignment Scores:  
 Pred. No.: 2.3e-51 Length: 407  
 Score: 633.00 Matches: 125  
 Percent Similarity: 94.07% Conservativity: 2  
 Best Local Similarity: 92.59% Mismatches: 8  
 Query Match: 85.54% Indels: 0  
 DB: 21 Gaps: 0

US-08-836-455-2 (1-145) x AAZ49548 (1-407)

Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
 Db 1 ATGAGGGCCCTGCTCAGATTGTTGGCTTCTTGTGCTCTGTTCCAGGTACCAGATGT 60  
 Qy 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
 Db 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCCCTCTCTGGGAGAAAGATCAGT 120  
 Qy 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeuGlnGluPro 60  
 Db 121 CTCACTTGTGGGCAAGTACGACATGTTGGTAAATATACCTGCTTCAACAGGAACCA 180  
 Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
 Db 181 GATGGAACCTTTAAACGCCCTGATCTACGCCACATCCAGTTAGATTCTGGTGTCCCAAG 240  
 Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
 Db 241 AGGTTCACTGGCAGTAGTCTGGTTCAGATTATCTCTCACCATCAGACGCTTGAGTCT 300  
 Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
 Db 301 GAGATTTTCTAGACTATTACTGTCTACAATATGCTAGTTCTCCGTACACGTTCCGAGGG 360  
 Qy 121 GlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrVal 135  
 Db 361 GGGACCAAGCTGGCAATAAACGGCTGATGCTGCACCACTGTA 405

RESULT 6

AAAT77851  
 ID AAT77851 standard; CDNA; 387 BP.

XX AC AAT77851;

XX 03-NOV-1997 (first entry)

XX Murine anti-human class II monoclonal antibody 44H104 VL chain cDNA.

XX Antibody: light chain; variable region; hybridoma cell line 44H104;  
 KW immune response; enhance; stimulate; vaccine; immunodiagnosis;  
 KW antigen delivery; ss.

XX OS Mus musculus.

XX Key Location/Qualifiers

XX CDS 1..387

XX /\*tag= a

XX /note= "Encodes 44H104 light chain variable region,

FT including secretion signal; termination  
 FT codon not given"

PN WO9640941-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-CA00400.

XX 07-JUN-1995; 95US-0483576.

XX (CONN-) CONNAUGHT LAB LTD.

XX Anand NN, Barber BH, Caterini JE, Cates GC, Klein MH;

XX WPI; 1997-077271/07.

XX P-PSDB; AAW22537.

XX Recombinant conjugate antibody mol., modified for delivering an  
 PT antigen - elicits enhanced immune response without the use of  
 PT adjuvant to generate antibodies which are useful in vaccines or  
 PT immuno:diagnosis

XX Example 1; Fig 1A; 64pp; English.

XX Novel recombinant conjugate antibody molecules comprise a monoclonal  
 CC antibody specific for a surface structure of antigen presenting  
 CC cells (APC), genetically modified to contain at least one antigen  
 CC exclusively at one or more preselected sites. The conjugate is capable  
 CC of delivering the antigen to APC and eliciting an immune response to  
 CC the antigen. The new conjugates are useful as vaccines and are able  
 CC to elicit an enhanced immune response without the use of an adjuvant.  
 CC In a specific example, a conjugate was constructed using the murine  
 CC anti-human class II monoclonal antibody secreted by hybridoma  
 CC 44H104. The peptide CLTB36 was chosen as antigen; it consists of  
 CC a tandemly linked T and B cell epitope derived from HIV MN strain.  
 CC The present sequence encodes the light chain variable region which  
 CC was PCR amplified from 44H104 and used in the preparation of a  
 CC conjugate with antigen CLTB36.

XX Sequence 387 BP; 90 A; 95 C; 90 G; 112 T; 0 other;

Alignment Scores:

Pred. No.: 2.39e-41 Length: 387  
 Score: 527.00 Matches: 103  
 Percent Similarity: 86.61% Conservativity: 7  
 Best Local Similarity: 81.10% Mismatches: 17  
 Query Match: 71.22% Indels: 0  
 DB: 18 Gaps: 0

US-08-836-455-2 (1-145) x AAT77851 (1-387)

Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
 Db 7 ATGAGGGTCTCTGCTCAGTTTGTGCTTCTGTTTCCAGGTACCAGATGT 66  
 Qy 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
 Db 67 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCCCTCTCTGGGACAAAGATCAGT 126  
 Qy 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeuGlnGluPro 60  
 Db 127 CTCACCTTGTGGGCAAGTACAGAAATAGTGGTTACTTAACCTGGCTTCCAGCAGAACCA 186  
 Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
 Db 187 GATGGAACCTATTAAACGCCCTGCTACGCCGCTCCTATTAGATTCTGGTGTCCCAAAA 246  
 Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
 Db 247 AGGTTCACTGGCAGTAGTCTGGTTCAGATTATCTCTCACCATCAGCAGCCTTGAGTCT 306  
 Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120

Db 307 GAGAGTTTGCAGACTATTACTGCTACAATATACTAATTATCGCTCAGCTTCGGTGCT 366  
 QY 121 GlyThrLysLeuGluLeuLys 127  
 Db 367 GGGACCAAGCTGGAGCTGAAA 387

## RESULT 7

AAD32138

ID AAD32138 standard; DNA; 387 BP.

XX AC AAD32138;

XX DT 18-JUN-2002 (first entry)

XX DE Murine 44H104 mab variable light chain (VL) DNA.

XX KW Murine; mab; light chain; VL; conjugate antibody; antigen delivery;

XX KW immune system; vaccine; detecting agent; antibacterial; gene; ds.

XX OS Mus sp.

XX FH Key Location/Qualifiers

XX FT 1..387

XX FT /product= "Murine 44H104 mab VL"

XX FT /transl\_except= (pos:277..279, aa:Thr)

XX FT /note= "This translation exception occurs while

XX FT decoding for murine mab VL alternative version

XX FT (AAE20204); CDS does not include stop codon"

XX FT /partial

XX PN US2002025315-A1.

XX PD 28-FEB-2002.

XX PF 14-JAN-1998; 98US-0007093.

XX PR 14-JAN-1998; 98US-0007093.

XX PA (ANAN/) ANAND N N.

XX PA (BARB/) BARBER B H.

XX PA (CATE/) CATES G A.

XX PA (CATE/) CATERINI J E.

XX PA (KLEI/) KLEIN M H.

XX PI Anand NN, Barber BH, Cates GA, Caterini JE, Klein MH;

XX DR WPI; 2002-267519/31.

XX DR P-PSDB; AAE20200, AAE20204.

XX PT Novel recombinant conjugate antibody, useful as a vaccine against

XX PT pathogens having a specific antigen, comprises a monoclonal antibody

XX PT specific for an antigen presenting cell surface structure -

XX PS Example 1; Fig 1A; 28pp; English.

XX CC The invention relates to a recombinant conjugate antibody, comprising

XX CC a monoclonal antibody specific for a surface structure of antigen

XX CC presenting cells genetically modified to contain an antigen moiety

XX CC for the purpose of delivery of the antigen moiety to antigen-

XX CC presenting cells of the immune system. The conjugate antibody is

XX CC formulated as a vaccine to protect a host against a disease caused

XX CC by a pathogen expressing the antigen. The antibody is useful as a

XX CC detecting agent. The present sequence is murine 44H104 mab variable

XX CC light chain (VL) DNA.

XX SQ Sequence 387 BP; 90 A; 95 C; 90 G; 112 T; 0 other;

## Alignment Scores:

Pred. No.: 2.39e-41

Score: 527.00

Percent Similarity: 86.61%

Length: 387

Matches: 103

Conservative: 7

Best Local Similarity: 81.10% Mismatches: 17  
 Query Match: 71.22% Indels: 0  
 DB: 24 Gaps: 0

US-08-836-455-2 (1-145) x AAD32138 (1-387)

QY 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20

Db 7 ATGAGGGTCTGCTCAGTTCACGTTTTTTGGCTTCTGTGCTCTGTGTTCCAGGTACCATGATGT 66

QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40

Db 67 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGACAAAGAGTCAGT 126

QY 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGlnPro 60

Db 127 CTCACITGTGCGGCAAGTCAGGAATAATAGTGGTTACTTAACCTGGCTTCAGCAACCA 186

QY 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80

Db 187 GATGGAACATATTAACGCTGTCTAGCGCGCTCCACTTTAGATTCTGGTGTCCCAAAA 246

QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100

Db 247 AGGTTCAGTGGCAGTAGGCTGGGTTCAGATTATCTTCACCATCAGCAGCCTTGAGTCT 306

QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120

Db 307 GAAGATTTTCAGAGCATATTACTGCTACAATATACTAATAATTCGCTCAGTTCGGTGCT 366

QY 121 GlyThrLysLeuGluLeuLys 127

Db 367 GGGACCAAGCTGGAGCTGAAA 387

## RESULT 8

AAQ46084

ID AAQ46084 standard; cDNA; 739 BP.

XX AC AAQ46084;

XX DT 25-MAR-2003 (updated)

XX DT 07-FEB-1994 (first entry)

XX DE Sequence encoding 520C9 sFv protein.

XX KW Single chain Fv polypeptide; VH-VL heterodimer; immunoglobulin; Ig;

XX KW biosynthetic single polypeptide chain binding site; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT CDS 1..739

XX FT /\*tag= a

XX PN WO9316185-A2.

XX PD 19-AUG-1993.

XX PF 05-FEB-1993; 93WO-US01055.

XX PR 06-FEB-1992; 92US-0831967.

XX PA (CETU ) CETUS ONCOLOGY CORP.

XX PA (CREA-) CREATIVE BIOMOLECULES INC.

XX PI Houston LL, Huston JS, Oppermann H, Ring DB;

XX DR WPI; 1993-272889/34.

XX DR P-PSDB; AAR39569.

XX PT New single chain Fv polypeptide binding to C-erbB-2 tumour

XX PT antigen - for imaging or treating breast or ovarian cancer etc.

PS Disclosure; pages 60-61; 87pp; English.

XX c-erbB-2 refers to a protein antigen expressed on the surface of

CC tumour cells, such as breast and ovarian tumour cells, which is an

CC approx. 200,000 mol. wt. acidic glycoprotein having an isoelectric

CC pt. of about 5.3 (see AAQ46083, AAR39568). A single chain Fv (sfv)

CC is a covalently linked VH-VL heterodimer which is expressed from

CC a gene fusion including VH- and VL- encoding genes connected by

CC a peptide-encoding linker. Such linker sequences are set forth in

CC AA residues 116-135 in AAR39569, which includes part of the 16 AA

CC linker sequences in AAR39572. Using AAQ46084 for the 520C9 monoclonal

CC antibody, a single chain polypeptide can be produced having a

CC binding affinity for a c-erbB-2 related antigen.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 739 BP; 188 A; 177 C; 181 G; 193 T; 0 other;

Alignment Scores:

Pred. No.: 3.95e-39 Length: 739

Score: 507.00 Matches: 101

Percent Similarity: 93.64% Conservative: 2

Best Local Similarity: 91.82% Mismatches: 7

Query Match: 68.51% Indels: 0

DB: 14 Gaps: 0

US-08-836-455-2 (1-145) x AAQ46084 (1-739)

QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40

DB 400 GATATCCAGATGACCCAGTCTCCATCTCTTATCTGCCCTCTGGGAGAAAGAGTCAGT 459

QY 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeuGlnGluPro 60

DB 460 CTCACCTTGCGGCAAGTCAGGACATTTGGTAATAGCTTAACCTGGCTTCAGCAGGAACCA 519

QY 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80

DB 520 GATGGAACTATTAAACGCCCTGATCTACGCCACATCCAGTTAGATTCTGGTGTGCCCAAA 579

QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100

DB 580 AGCTTCAGTGGCAGTGGTGGTCAGATTATCTCTCACCATCAGTACCTTGAGTCT 639

QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120

DB 640 GAAGATTTTGTAGTCTATTACTGCTACAAATATGCTATTTTCCGTACACGTTCCGGAGGG 699

QY 121 GlyThrLysLeuGluIleLysArgAlaAsp 130

DB 700 GGGACCAACCTGGAAATAAACGGGCTGAT 729

RESULT 9

AAT36880

ID AAT36880 standard; cDNA; 739 BP.

XX

AC AAT36880;

XX

DT 25-MAR-2003 (updated)

DT 29-OCT-1996 (first entry)

XX

XX 520C9 anti-c-erbB-2 two single chain Fv construct.

XX

KW 520C9; anti-c-erbB-2 monoclonal antibody; single chain Fv; sfv;

KW construct; polypeptide linker; c-terminal amino acid sequence;

KW in vivo imaging; drug targeting experiment; homodimer;

KW increased; binding avidity; tissue retention time; ss.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FF CDS 1..732

FF /\*tag= a

FF /note= "START codon absent"

XX US5534254-A.

PN

XX

XX 09-JUL-1996.

XX

XX 07-OCT-1993; 93US-0133804.

PF

XX 07-OCT-1993; 93US-0133804.

PR

XX 06-FEB-1992; 92US-0831967.

PR

XX (CHIR ) CHIRON CORP.

PA (CREA-) CREATIVE BIOMOLECULES INC.

XX

XX Houston LL, Huston JS, Oppermann H, Ring DB;

PI

XX WPI; 1996-333194/33.

DR

XX P-PSDB; AAW02280.

DR

XX

XX Compens. contg. antigen-targetting antibody fragment constructs -

PT comprising dimer of single-chain Fv fragments

PT

XX Example 1; Columns 33-36; 30pp; English.

PS

XX

XX Variable heavy (VH) and variable light (VL) genes were cloned from

CC a 520C9 hybridoma cDNA library, using probes directed toward the

CC antibody constant and joining regions. A two single chain Fv (sfv)

CC gene was constructed by connecting the VH and VL genes with a

CC Ser rich polypeptide linker. The resulting 520C9 two sfv gene, the

CC present sequence, was inserted into an expression vector, by the

CC transformed into E. coli, and protein expression induced, by the

CC addn. of IPTG to the culture medium.

CC A compsn. comprising a carrier and the 2 sfv protein prod. can be

CC used for in vivo imaging, and drug targeting experiments. The

CC 2 sfv protein prod. is a homodimer, in which both fragments target

CC the same antigen, therefore giving greater binding avidity and

CC longer tissue retention times, compared to individual sfv protein

CC prod. fragments.

CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 739 BP; 188 A; 177 C; 181 G; 193 T; 0 other;

Alignment Scores:

Pred. No.: 3.95e-39 Length: 739

Score: 507.00 Matches: 101

Percent Similarity: 93.64% Conservative: 2

Best Local Similarity: 91.82% Mismatches: 7

Query Match: 68.51% Indels: 0

DB: 17 Gaps: 0

US-08-836-455-2 (1-145) x AAT36880 (1-739)

QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40

DB 400 GATATCCAGATGACCCAGTCTCCATCTCTTATCTGCCCTCTGGGAGAAAGAGTCAGT 459

QY 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeuGlnGluPro 60

DB 460 CTCACCTTGCGGCAAGTCAGGACATTTGGTAATAGCTTAACCTGGCTTCAGCAGGAACCA 519

QY 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80

DB 520 GATGGAACTATTAAACGCCCTGATCTACGCCACATCCAGTTAGATTCTGGTGTGCCCAAA 579

QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100

DB 580 AGCTTCAGTGGCAGTGGTGGTCAGATTATCTCTCACCATCAGTACCTTGAGTCT 639

QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120

DB 640 GAAGATTTTGTAGTCTATTACTGCTACAAATATGCTATTTTCCGTACACGTTCCGGAGGG 699

QY 121 GlyThrLysLeuGluIleLysArgAlaAsp 130

DB 700 GGGACCAACCTGGAAATAAACGGGCTGAT 729

```

Db      700 GGGACCAACCTGGAAATAAAACGGGCTGAT 729
RESULT 10
AAV21798
ID      AAV21798 standard; cDNA; 739 BP.
XX
AC      AAV21798;
XX
XX      16-JUL-1998 (first entry)
XX
DE      520C9 anti-c-erbB-2 sfv' dimeric construct gene sequence.
XX
XX      Antigen imaging; single chain Fv; sfv; linker; dimeric; cancer;
KW      c-erbB-2; tumour; diagnosis; ss.
XX
XX      Synthetic.
OS      Mus sp.
XX
XX      Location/Qualifiers
FH      Key
CDS     1..732
FT      /*tag= a
FT      /product= "520C9 sfv' polypeptide"
XX
XX      US5753204-A.
XX
XX      19-MAY-1998.
XX
XX      05-JUN-1995; 95US-0461838.
XX
XX      07-OCT-1993; 93US-0133804.
PR      06-FEB-1992; 92US-0831967.
PR      05-JUN-1995; 95US-0461838.
XX
XX      (CHIR ) CHIRON CORP.
PA      (CREA-) CREATIVE BIOMOLECULES INC.
XX
XX      Houston LL, Huston JS, Oppermann H, Ring DB;
XX
XX      WPI; 1998-311318/27.
DR      P-PSDB; AAW53170.
XX
XX      Imaging of antigens in vivo - using dimers of single-chain antibody
XX      Fv fragments
XX
XX      Example 1; Columns 33-36; 30pp; English.
XX
XX      This is the nucleotide sequence of a 520C9 sfv' (single chain Fv)
XX      construct. This was constructed by connecting the Vh and Vl genes with
XX      a DNA sequence encoding a Serine rich linker. 520C9 is a monoclonal
XX      antibody useful in targeting c-erbB-2 antigen. This dimeric construct
XX      can be used in the methods of invention of imaging a preselected antigen
XX      expressed in a mammal. The methods are used in magnetic resonance imaging
XX      of c-erbB-2 or related antigens in cancer diagnosis. The biosynthetic
XX      constructs have enhanced properties as in vivo targeting agents in
XX      comparison with intact monoclonal antibodies or their Fab fragments. The
XX      dimeric constructs permit the in vivo targeting of an epitope on an
XX      antigen with greater apparent avidity, including greater tumour
XX      specificity, tumour localisation and tumour retention properties than
XX      that of the Fab fragment having the same CDRS as the construct.
XX
XX      Sequence 739 BP; 188 A; 177 C; 181 G; 193 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 3.95e-39 Length: 739
Score: 507.00 Matches: 101
Percent Similarity: 93.64% Conservative: 2
Best Local Similarity: 91.82% Mismatches: 7
Query Match: 68.51% Indels: 0
DB: 19 Gaps: 0

us-08-836-455-2 (1-145) x AAV21798 (1-739)
Qy      21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40

```

```

|||||
400 GATATCCAGATGACCCAGTCTCCATCTCTTATCTGCCTCTCTGGGAGAAAGAGTCAGT 459
Qy      41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGlnPro 60
|||||
460 CTCACITGTCGGGCAAGTCAGGACATGGTAATAGCTTTAACCTGGCTTCAGCAGGAACCA 519
Qy      61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80
|||||
520 GATGGAACTATTAAACGCCTGATCTACGCCACATCCAGTTAGATTCTGGTGTCCCAAA 579
Qy      81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGlySer 100
|||||
580 AGGTTCAGTGGCAGTCTGGCTCAGATTATTTCTCTCACCATCATAGCTTGAGTCT 639
Qy      101 GluAspPheValAlaLayTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120
|||||
640 GAAGATTTTGTAGTCTATTACTGTCTACAATATGCTATTTTTCGCTACACGCTCGGAGG 699
Qy      121 GlyThrLysLeuGluIleLysArgAlaAsp 130
|||||
700 GGGACCAACCTGGAAATAAAACGGGCTGAT 729

RESULT 11
AAV63399
ID      AAV63399 standard; cDNA; 739 BP.
XX
XX      AAV63399;
XX
XX      28-JAN-1999 (first entry)
XX
XX      520C9 sfv DNA sequence.
XX
XX      520C9 sfv; antigen; tumour cell; antibody 520C9;
KW      targeted delivery; antigen-expressing cell; ss.
XX
XX      Synthetic.
OS
XX
XX      Key
FH      Location/Qualifiers
CDS     1..732
FT      /*tag= a
FT      /product= 520C9 sfv
XX
XX      US5837846-A.
XX
XX      17-NOV-1998.
XX
XX      05-JUN-1995; 95US-0461386.
XX
XX      07-OCT-1993; 93US-0133804.
PR      06-FEB-1992; 92US-0831967.
PR      05-JUN-1995; 95US-0461386.
XX
XX      (CHIR ) CHIRON CORP.
PA      (CREA-) CREATIVE BIOMOLECULES INC.
XX
XX      Houston LL, Huston JS, Oppermann H, Ring DB;
XX
XX      WPI; 1999-023541/02.
DR      P-PSDB; AAW80424.
XX
XX      Nucleic acid encoding single-chain Fv fragment specific for antigens
XX      - and having C-terminal tail for crosslinking to form dimer with
XX      improved pharmacokinetic properties, used to deliver drugs and
XX      imaging agents, especially to tumours
XX
XX      Example 1; Columns 33-36; 29pp; English.
XX
XX      The present sequence encodes an antibody 520C9 sfv. Variable heavy
XX      and light sequences of antibody 7520C9 are connected, together with a
XX      serine linker, to produce the present single chain Fv gene. The present
XX      sequence exemplifies the invention. Dimers of the single chain Fv are
XX      used for targeted delivery of drugs or imaging agents (e.g. cytotoxins,

```

CC prodrugs or 99m-technetium) to antigen-expressing cells, particularly  
 XX for treatment or diagnosis of tumours (especially of ovary or breast).  
 SQ Sequence 739 BP; 188 A; 177 C; 181 G; 193 T; 0 other;

Alignment Scores:  
 Pred. NO.: 3.95e-39 Length: 739  
 Score: 507.00 Matches: 101  
 Percent Similarity: 93.64% Conservative: 2  
 Best Local Similarity: 91.82% Mismatches: 7  
 Query Match: 68.51% Indels: 0  
 DB: 20 Gaps: 0

US-08-836-455-2 (1-145) x AAV63399 (1-739)

Qy 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
 Db 400 GATATCCAGATGACCCAGTCCTCTTATCTGCTCTCTGGGAGAAAGAGTCAGT 459  
 Qy 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuHisTrpLeuGlnGlnGluPro 60  
 Db 460 CTCACCTTGTCTGGGCAAGTCAGGACATTTGGTAATAGCTTAACCTGGCTTCAGCAGGAACCA 519  
 Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
 Db 520 GATGGAACTATTAAACGCCGTGATCTACGCCACATCCAGTTTAGATTCTGGTGTCCCAAA 579  
 Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
 Db 580 AGGTTTCAGTGGCAGTCGGTGGTCAGATTATCTCTCACCATCAGTAGCCTTGAGTCT 639  
 Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
 Db 640 GAAGATTTTGTAGTCTATTACTGTCTACAATATGCTATTTTCCGTACACGTTTCGGAGGG 699  
 Qy 121 GlyThrLysLeuGluIleLysArgAlaasp 130  
 Db 700 GGGACCAACCTCGGAATAAAACGGGCTGAT 729

RESULT 12  
 AAZ49534  
 ID AAZ49534 standard; cDNA: 324 BP.  
 AC AAZ49534;  
 XX  
 DT 04-APR-2000 (first entry)  
 XX  
 DE Mouse anti-IL-18 antibody VL encoding cDNA from hybridoma #125-2H.  
 XX  
 KW Anti-Interleukin-18 antibody; IL-18; light chain variable region; VL;  
 KW hybridoma #125-2H; mouse; monoclonal antibody #125-2HmAb; Interleukin-18;  
 KW antiinflammatory; immunosuppressive; leucocytopenic; antialgic;  
 KW antipyretic; antiallergic; hepatotropic; treatment; autoimmune disease;  
 KW immunopathy; inflammatory disorder; immunoreaction; ss.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT mat\_peptide 1..324  
 FT /\*tag= a  
 FT /label= Anti-IL-18\_antibody\_light\_chain\_variable\_region  
 XX  
 EP974600-A2.  
 XX  
 PD 26-JAN-2000.  
 XX  
 PF 24-JUN-1999; 99EP-0304977.  
 XX  
 PR 24-JUN-1998; 98JP-0177580.  
 PR 12-OCT-1998; 98JP-0289044.  
 PR 22-DEC-1998; 98JP-0365023.  
 XX  
 PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.

XX Nishida Y, Okura T, Tanimoto T, Kurimoto M;  
 XX WPI: 2000-118341/11.  
 DR P-PSDB: AAY44587.  
 XX  
 PT New artificially produced peptide for neutralizing biological activity  
 PT of interleukin-18, useful for treating and preventing immunopathies,  
 XX inflammatory disorders and autoimmune diseases -  
 PS Claim 11; Page 21; 36pp; English.  
 XX  
 CC The present cDNA sequence derived from hybridoma #125-2H, encodes mouse  
 CC anti-interleukin-18 antibody light chain variable region (VL). It can be  
 CC used in the production of recombinant monoclonal antibody #125-2HmAb,  
 CC which is capable of neutralising biological activities of interleukin-18.  
 CC The antibody has antiinflammatory; immunosuppressive; leucocytopenic;  
 CC antialgic, antipyretic, antiallergic and hepatotropic activity and can be  
 CC used for prevention and treatment of autoimmune diseases, immunopathies  
 CC and inflammatory disorders caused by excessive immunoreaction.  
 XX  
 SQ Sequence 324 BP; 83 A; 78 C; 76 G; 87 T; 0 other;

Alignment Scores:  
 Pred. NO.: 1.9e-39 Length: 324  
 Score: 506.00 Matches: 100  
 Percent Similarity: 94.44% Conservative: 2  
 Best Local Similarity: 92.59% Mismatches: 6  
 Query Match: 68.38% Indels: 0  
 DB: 21 Gaps: 0

US-08-836-455-2 (1-145) x AAZ49534 (1-324)

Qy 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
 Db 1 GACATCCAGATGACCCAGTCCTCTTATCTGCTCTCTGGGAGAAAGAGTCAGT 60  
 Qy 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuHisTrpLeuGlnGlnGluPro 60  
 Db 61 CTCACCTTGTCTGGGCAAGTCAGGACATTTGGTAATAGCTTAACCTGGCTTCAGCAGGAACCA 120  
 Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
 Db 121 GATGGAACTTTTAAACGCCGTGATCTACGCCACATCCAGTTTAGATTCTGGTGTCCCAAG 180  
 Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
 Db 181 AGGTTTCAGTGGCAGTCAGTCTGGTTCAGATTATCTCTCACCATCAGCAGCCTTGAGTCT 240  
 Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
 Db 241 GAAGATTTTGTAGACTATTACTGTCTACAATATGCTAGTTCTCGGTACACGTTTCGGAGGG 300  
 Qy 121 GlyThrLysLeuGluIleLysArg 128  
 Db 301 GGGACCAAGCTGGCAATAAAACGG 324

RESULT 13  
 AAS97136  
 ID AAS97136 standard; DNA: 756 BP.  
 XX  
 AC AAS97136;  
 XX  
 DT 26-FEB-2002 (first entry)  
 XX  
 DE P4-3 single chain Fv DNA.  
 XX  
 KW Human; NKG2D; NKG2D receptor complex; cancer; infectious disease; tumour;  
 KW autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;  
 KW intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;  
 KW prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;  
 KW sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; ds;  
 KW helminth; cytostatic; antimicrobial; immunomodulatory; P4-2; P4-3; P4-14;







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 30, 2003, 22:11:56 ; Search time 192.198 seconds  
(without alignments)  
1730.853 Million cell updates/sec

Title: US-08-836-455-2

Perfect score: 740

Sequence: 1 MGAPQILGFLLLFPQTRC.....IKRADAAPTIVSIPPPSKLG 145

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US08836455/runat\_29082003\_132902\_22302/app\_query.fasta\_1.654  
-DB=PublishedApplications\_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosom62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US08836455 -CGN1\_1.271 -runat\_29082003\_132902\_22302  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PublishedApplications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/PCTUS\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	740	100.0	435	9	US-09-861-294-1 Sequence 1, Appli

2	740	100.0	435	12	US-10-367-506-1 Sequence 1, Appli
3	633	85.5	407	10	US-09-924-099-27 Sequence 27, Appli
4	616	83.2	402	8	US-08-779-784-5 Sequence 5, Appli
5	541	73.1	390	12	US-10-268-883-10 Sequence 10, Appli
6	532	71.9	351	8	US-08-779-784-16 Sequence 16, Appli
7	532	71.9	351	8	US-08-779-784-17 Sequence 17, Appli
8	527	71.2	387	9	US-09-007-093-1 Sequence 1, Appli
9	507	68.5	739	10	US-09-887-853-5 Sequence 5, Appli
10	506	68.4	324	10	US-09-924-099-11 Sequence 11, Appli
11	502	67.8	711	10	US-09-924-099-19 Sequence 19, Appli
12	502	67.8	729	10	US-09-924-099-20 Sequence 20, Appli
13	484.5	65.5	737	10	US-09-919-344-7 Sequence 7, Appli
14	481	65.0	714	14	US-10-153-382-18 Sequence 18, Appli
15	473	63.9	974	10	US-09-859-053-29 Sequence 29, Appli
16	471	63.6	447	9	US-09-797-481-1 Sequence 1, Appli
17	471	63.6	447	9	US-09-844-736-3 Sequence 3, Appli
18	471	63.6	447	14	US-10-162-396-3 Sequence 3, Appli
19	469	63.4	490	11	US-09-918-995-37859 Sequence 37859, A
20	467	63.1	447	9	US-09-861-294-19 Sequence 19, Appli
21	467	63.1	447	12	US-10-367-506-19 Sequence 19, Appli
22	464	62.7	634	14	US-10-158-646-55 Sequence 55, Appli
23	464	62.7	728	10	US-09-844-684-15 Sequence 15, Appli
24	464	62.7	728	14	US-10-040-244-15 Sequence 15, Appli
25	462	62.4	716	10	US-09-844-684-13 Sequence 13, Appli
26	462	62.4	716	14	US-10-040-244-13 Sequence 13, Appli
27	462	62.4	729	14	US-10-216-484-125 Sequence 125, App
28	461	62.3	393	8	US-08-779-784-20 Sequence 20, Appli
29	455.5	61.6	928	14	US-10-221-945-5 Sequence 5, Appli
30	447	60.4	514	14	US-10-066-543-2025 Sequence 2025, Ap
c 31	447	60.4	537	14	US-10-066-543-186 Sequence 186, App
32	445	60.1	705	9	US-09-740-002-16 Sequence 16, Appli
33	444	60.0	840	9	US-09-766-543-11 Sequence 11, Appli
c 34	442	59.7	520	10	US-09-878-178-1210 Sequence 1210, Ap
c 35	442	59.7	520	13	US-10-046-935-1210 Sequence 1210, Ap
c 36	442	59.7	520	14	US-10-146-502-1210 Sequence 1210, Ap
37	439	59.3	381	12	US-10-268-883-4 Sequence 4, Appli
38	439	59.3	819	14	US-10-158-646-65 Sequence 65, Appli
39	438	59.2	698	10	US-09-844-684-11 Sequence 11, Appli
40	438	59.2	698	14	US-10-040-244-11 Sequence 11, Appli
41	438	59.2	711	14	US-10-006-593-70 Sequence 70, Appli
42	436.5	59.0	494	10	US-09-878-178-1811 Sequence 1811, Ap
43	436.5	59.0	494	13	US-10-046-935-1811 Sequence 1811, Ap
44	436.5	59.0	494	14	US-10-146-502-1811 Sequence 1811, Ap
45	436	58.9	408	9	US-09-764-304-9 Sequence 9, Appli

ALIGNMENTS

RESULT 1  
US-09-861-294-1  
; Sequence 1, Application US/09861294  
; Patent No. US20020098190A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TUMORS BEARING HMFG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/09/861,294  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 435  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(435)

; NAME/KEY: sig\_peptide  
; LOCATION: (1)...(60)  
; NAME/KEY: mat\_peptide  
; LOCATION: (61)...(435)  
US-09-861-294-1

## Alignment Scores:

Pred. No.: 1.48e-84 Length: 435  
Score: 740.00 Matches: 145  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-08-836-455-2 (1-145) x US-09-861-294-1 (1-435)

QY 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuLeuPheProGlyThrArgCys 20  
DB 1 ATGGGGCCCTGCTCAGATTCTTGGGTCTTGTGCTCTTGTTCAGGTACCAGATGT 60  
QY 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
DB 61 GACATCCAGATGACCCAGCTCCATCTCTTATCTGCTCTCTGGGACAAAGATCAGT 120  
QY 41 LeuThrCysArgAlaSerGlnAspileGlyIleAsnLeuHisTrpLeuGlnGlnPro 60  
DB 121 CTCACCTTGTGGGCAAGTCAGACATGCTGTTAACTTACATTTGCTTCCAGCAGCA 180  
QY 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerLeuGlySerGlyValProLys 80  
DB 181 GATGGAACATTAACGGCTGATCTACGCCATCCAGTATTCTCCACCATCAGCAGCTT 240  
QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
DB 241 AGGTTACGTGGCAGTAGCTCTGGGTACAGATTATCTCTCACCATCAGCAGCTT 300  
QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
DB 301 GAAGATTTGTAGCTATTACTGTCTACAATATGCTAGTTCTCCGATACACGTT 360  
QY 121 GlyThrLysLeuGluIleLysArgAlaAspAlaProThrValSerIlePheProPro 140  
DB 361 GGGACCAAGCTGGAATAAACGGGCTGATGCTGCACCACTGATATCTTCCACCA 420  
QY 141 SerSerLysLeuGly 145  
DB 421 TCCAGTAAGCTTGGG 435

## RESULT 2

US-10-367-506-1  
; Sequence 1, Application US/10367506  
; Publication No. US20030152575A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TITLE OF INVENTION: TUMORS BEARING HMEG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/10/367,506  
; CURRENT FILING DATE: 2003-02-13  
; PRIOR APPLICATION NUMBER: US/09/861,294  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 435  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:

; NAME/KEY: CDS  
; LOCATION: (1)...(435)  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: (1)...(60)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (61)...(435)  
US-10-367-506-1

## Alignment Scores:

Pred. No.: 1.48e-84 Length: 435  
Score: 740.00 Matches: 145  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-08-836-455-2 (1-145) x US-10-367-506-1 (1-435)

QY 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuLeuPheProGlyThrArgCys 20  
DB 1 ATGGGGCCCTGCTCAGATTCTTGGGTCTTGTGCTCTTGTTCAGGTACCAGATGT 60  
QY 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
DB 61 GACATCCAGATGACCCAGCTCCATCTCTTATCTGCTCTCTGGGACAAAGATCAGT 120  
QY 41 LeuThrCysArgAlaSerGlnAspileGlyIleAsnLeuHisTrpLeuGlnGlnPro 60  
DB 121 CTCACCTTGTGGGCAAGTCAGACATGCTGTTAACTTACATTTGCTTCCAGCAGCA 180  
QY 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerLeuGlySerGlyValProLys 80  
DB 181 GATGGAACATTAACGGCTGATCTACGCCATCCAGTATTAGTCTGTTGCCCA 240  
QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
DB 241 AGGTTACGTGGCAGTAGCTCTGGGTACAGATTATCTCTCACCATCAGCAGCTT 300  
QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
DB 301 GAAGATTTGTAGCTATTACTGTCTACAATATGCTAGTTCTCCGATACACGTT 360  
QY 121 GlyThrLysLeuGluIleLysArgAlaAspAlaProThrValSerIlePheProPro 140  
DB 361 GGGACCAAGCTGGAATAAACGGGCTGATGCTGCACCACTGATATCTTCCACCA 420  
QY 141 SerSerLysLeuGly 145  
DB 421 TCCAGTAAGCTTGGG 435

## RESULT 3

US-09-924-099-27  
; Sequence 27, Application US/09924099  
; Patent No. US20020128450A1  
; GENERAL INFORMATION:  
; APPLICANT: NISHIDA, Yoshihiro  
; APPLICANT: OKURA, Takanori  
; APPLICANT: TANIMOTO, Tadao  
; APPLICANT: KURIMOTO, Masashi  
; TITLE OF INVENTION: PEPTIDE  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/924,099  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 33  
; SEQ ID NO 27

```

; LENGTH: 407
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(407)
; NAME/KEY: sig peptide
; LOCATION: (1)...(60)
US-09-924-099-27

Alignment Scores:
Pred. No.: 4,79e-71 Length: 407
Score: 633.00 Matches: 125
Percent Similarity: 94.07% Conservative: 2
Best Local Similarity: 92.59% Mismatches: 8
Query Match: 85.54% Indels: 0
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x US-09-924-099-27 (1-407)

Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20
Db 1 ATGAGGCCCTGCTCAGATTTTGGCTTCTTGTGCTCTGTTCAGGTACCAGATGT 60
Qy 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlnArgValSer 40
Db 61 GACATCCAGATGACCCAGTCTCCATCCTCTTATCTGCCTCTCTGGGAGAAAGAGTCAGT 120
Qy 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeuGlnGluPro 60
Db 121 CTCACTTGTGGGCAAGTCAGACATTTGGTAGTAATAATATATCTTCAACAGGAACCA 180
Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80
Db 181 GATGGAACCTTTAAACGCCCTGATCTAGCCACATCCAGTTAGATTCTGGTGTCCCAAG 240
Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100
Db 241 AGGTTAGTGGCAGTAGGTCTGGGTGAGATTTCTCTCACCATCAGCAGCCTTGAGTCT 300
Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120
Db 301 GAAGATTTCTAGACATTACTGTCTACATATGCTAGTTCTCCGTACACGTTCCGAGGG 360
Qy 121 GlyThrLysLeuGluIleLysArgAlaAspAlaProThrVal 135
Db 361 GGGACCAAGCTGGCAATAAAACGGCTGATGCTGCACCAACTGTA 405

RESULT 4
US-08-779-784-5
; Sequence 5, Application US/08779784
; Publication No. US20020164325A1
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Moses
; APPLICANT: Miller, David J.
; APPLICANT: Asakura, Kunihiko
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM
; TITLE OF INVENTION: REMYELINATION USING MONOCLONAL AUTOANTIBODIES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/779,784
; FILING DATE: 07-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/692,084
; FILING DATE: 08-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/236,520
; FILING DATE: 29-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; US-08-779-784-5

Alignment Scores:
Pred. No.: 6,69e-69 Length: 402
Score: 616.00 Matches: 122
Percent Similarity: 95.45% Conservative: 4
Best Local Similarity: 92.42% Mismatches: 6
Query Match: 83.24% Indels: 0
DB: 8 Gaps: 0

US-08-836-455-2 (1-145) x US-08-779-784-5 (1-402)

Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20
Db 7 ATGAGGCCCTGCTCAGACAGATTTTGGCTTCTTGTGCTCTGTTCAGGTACCAGATGT 66
Qy 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlnArgValSer 40
Db 67 GACATCCAGATGACCCAGTCTCCATCCTCTTATCTGCCTCTCTGGGAGAAAGAGTCAGT 126
Qy 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeuGlnGluPro 60
Db 127 CTCACTTGTGGGCAAGTCAGACATTTGGTAGTAGCTTAACTGGCTTCAGCAGGAACCA 186
Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80
Db 187 GATGGAACCTTTAAACGCCCTGATCTAGCCACATCCAGTTAGATTCTGTGTGCCCAAA 246
Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100
Db 247 AGGTTAGTGGCAGTAGGTCTGGGTGAGATTTCTCTCACCATCAGCAGCCTTGAGTCT 306
Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120
Db 307 GAAGATTTCTAGACTATTACTGTCTACATATGCTAGTTCTCCGTACACGTTCCGAGGG 366
Qy 121 GlyThrLysLeuGluIleLysArgAlaAspAla 132
Db 367 GGGACCAAGCTGGAAATAAAACGGCTGATGCTTCA 402

RESULT 5
US-10-268-883-10
; Sequence 10, Application US/10268883
; Publication No. US20030138862A1
; GENERAL INFORMATION:
; APPLICANT: Tso, J. Yun
; APPLICANT: Green, Jennifer Macphate

```





QY 61 AspGlyThrLeuLysArgLeuLeuTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
|||||  
Db 187 GATGCAACTATTAAACGCGTGGTCTACGCGGCTCCACTTTAGATTCTGGTGTCACAAA 246  
QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
|||||  
Db 247 AGTTTCAGTCGCGAGTGGTCTGGGTACAGATTATCTCCACCATCAGCAGCCTTGAGTCT 306  
QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
|||||  
Db 307 GAAGATTTTGCAGACTATTACTGTCTACAATATACTAATATATCCGCTCAGCTTCGGTCT 366  
QY 121 GlyThrLysLeuGluLeuLys 127  
|||||  
Db 367 GGGACCAAGCTGGAGCTGAAA 387

## RESULT 9

US-09-887-853-5  
; Sequence 5, Application US/09887853  
; Patent No. US20020168375A1  
; GENERAL INFORMATION:  
; APPLICANT: Huston, James S.  
; Oppermann, Hermann  
; Houston, L. L.  
; Ring, David B.  
; TITLE OF INVENTION: Biosynthetic Binding Proteins For  
; Imaging  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department  
; STREET: Exchange Place, 53 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/887,853  
FILING DATE: 21-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/133,804  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Kelley, Robin D.  
REGISTRATION NUMBER: 34,637  
REFERENCE/DOCKET NUMBER: 2054/22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-248-7477  
TELEFAX: 617-248-7100  
INFORMATION FOR SEQ ID NO: 5:

## SEQUENCE CHARACTERISTICS:

LENGTH: 739 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..729  
OTHER INFORMATION: /product= "520C9 sfv polypeptide  
sequence"  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:

## US-09-887-853-5

Alignment Scores:  
Pred. No.: 9,53e-55 Length: 739  
Score: 507.00 Matches: 101

Percent Similarity: 93.64% Conservative: 2  
Best Local Similarity: 91.82% Mismatches: 7  
Query Match: 68.51% Indels: 0  
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x US-09-887-853-5 (1-739)

QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
|||||  
Db 400 GATATCCAGATGACCCAGCTCTCCATCCTCTATCTGCTCTCTGCGGAGAAAGAGTCACT 459  
QY 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGlnPro 60  
|||||  
Db 460 CTCACCTTTCGCGCAAGTCAGGACATTTGGTAAATAGCTTTAACTTTCCTCAGCAGGAACA 519  
QY 61 AspGlyThrIleLysArgLeuLeuTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
|||||  
Db 520 GATGCAACTATTAAACGCGTGGTCTACGCGGCTCCAGTTAGATTCTGGTGTCACAAA 579  
QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
|||||  
Db 580 AGTTTCAGTCGCGAGTGGTCTGGGTACAGATTATCTCTCACCATCAGTAGCCTTGAGTCT 639  
QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
|||||  
Db 640 GAAGATTTTGTAGTCTATTACTGTCTACAATATGCTATTTTCCGTACACGTTCCGAGGG 699  
QY 121 GlyThrLysLeuGluLeuLysArgAlaAsp 130  
|||||  
Db 700 GGGACCAACCTGGAAATAAAACGGGCTGAT 729

## RESULT 10

US-09-924-099-11  
; Sequence 11, Application US/09924099  
; Patent No. US20020128450A1

## GENERAL INFORMATION:

; APPLICANT: NISHIDA, Yoshihiro  
; APPLICANT: OKURA, Takanori  
; APPLICANT: TANIMOTO, Tadao  
; APPLICANT: KURIMOTO, Masashi  
; TITLE OF INVENTION: PEPTIDE  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/924,099  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 33

; SEQ ID NO 11

; LENGTH: 324

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-924-099-11

## Alignment Scores:

Pred. No.: 4.25e-55 Length: 324  
Score: 506.00 Matches: 100  
Percent Similarity: 94.44% Conservative: 2  
Best Local Similarity: 92.59% Mismatches: 6  
Query Match: 68.38% Indels: 0  
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x US-09-924-099-11 (1-324)

QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
|||||  
Db 1 GACATCCAGATGACCCAGCTCTCCATCCTCTATCTGCTCTCTGCGGAGAAAGAGTCACT 60  
QY 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGlnPro 60  
|||||



Db 61 CTCACCTGTGGGCAAGTCAGGACATTTGGTAGTAAATATTACTGGCTTCAACAGGACCA 120  
QY 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
Db 121 GATGGAACTTTAAAGCGCTGATCTACGCCACATCCAGTTAGATTCTGGTGTCCCAAG 180  
QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGlySer 100  
Db 181 AGCTTCAGTGGCAGTAGTCTGGTGCAGATTATCTCTCACCACATCAGCAGCCCTGAGTCT 240  
QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
Db 241 GAAGATTTTGTAGACTATTACTCTCACAATATGCTAGTTCTCCGTACACGTTCCGGAGGG 300  
QY 121 GlyThrLysLeuGluIleLysArg 128  
Db 301 GGGACCAAGCTGGCAATAAAACGG 324  
RESULT 11  
US-09-924-099-19  
; Sequence 19, Application US/09924099  
; Patent No. US20020128450A1  
; GENERAL INFORMATION:  
; APPLICANT: NISHIDA, Yoshihiro  
; APPLICANT: OKURA, Takanori  
; APPLICANT: TANIMOTO, Tadao  
; APPLICANT: KURIMOTO, Masashi  
; TITLE OF INVENTION: PEPTIDE  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/924,099  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-06-23  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-10-12  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12  
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 33  
; SEQ ID NO 19  
; LENGTH: 711  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificial DNA to code for the amino acid sequence of  
; OTHER INFORMATION: SEQ  
; OTHER INFORMATION: ID NO:9  
US-09-924-099-19  
Alignment Scores:  
Pred. No.: 3.89e-54 Length: 711  
Score: 502.00 Matches: 100  
Percent Similarity: 91.89% Conservative: 2  
Best Local Similarity: 90.09% Mismatches: 9  
Query Match: 67.84% Indels: 0  
DB: 10 Gaps: 0  
US-08-836-455-2 (1-145) x US-09-924-099-19 (1-711)  
QY 17 GlyThrArgCysAspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGly 36  
Db 379 GGTGGCGGATCGGACATCCAGATGACCCAGTCTCCATCTCTCTATCTGCTCTCTGGGA 438  
QY 37 GlnArgValSerLeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeu 56  
Db 439 GAAGAGTCACTCTCACTTCTCGGGCAAGTCAGGACATTTAGTAAATATATCTGCTT 498  
QY 57 GlnGlnGluProAspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySer 76  
Db 499 CAACAGCAACAGATGGAACCTTTTAAACGCTGATCTACGCCACATCCAGTTTATGATCT 558  
QY 77 GlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSer 96  
Db 559 GGTGTCCCAAGAGGTTTCAGTGGCAGTAGGTCTGGTCCAGATTTATCTCTCACCATCAGC 618  
QY 97 SerLeuGluSerGluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyr 116  
Db 619 AGCTTCAGTCTGAAGATTTTGTAGACTATTACTGTCTACAATATGCTAGTCTCCGTAC 678  
QY 117 ThrPheGlyGlyThrLysLeuGluIleLys 127  
Db 679 ACGTTCGAGGGGGGACCAAGCTGGCAATAAAA 711

QY 97 SerLeuGluSerGluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyr 116  
Db 619 AGCTTCAGTCTGAAGATTTTGTAGACTATTACTGTCTACAATATGCTAGTCTCCGTAC 678  
QY 117 ThrPheGlyGlyThrLysLeuGluIleLys 127  
Db 679 ACGTTCGAGGGGGGACCAAGCTGGCAATAAAA 711  
RESULT 12  
US-09-924-099-20  
; Sequence 20, Application US/09924099  
; Patent No. US20020128450A1  
; GENERAL INFORMATION:  
; APPLICANT: NISHIDA, Yoshihiro  
; APPLICANT: OKURA, Takanori  
; APPLICANT: TANIMOTO, Tadao  
; APPLICANT: KURIMOTO, Masashi  
; TITLE OF INVENTION: PEPTIDE  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/924,099  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-06-23  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-10-12  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12  
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 33  
; SEQ ID NO 20  
; LENGTH: 729  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificial DNA to code for the amino acid sequence of  
; OTHER INFORMATION: SEQ  
; OTHER INFORMATION: ID NO:10  
US-09-924-099-20  
Alignment Scores:  
Pred. No.: 4.02e-54 Length: 729  
Score: 502.00 Matches: 100  
Percent Similarity: 91.89% Conservative: 2  
Best Local Similarity: 90.09% Mismatches: 9  
Query Match: 67.84% Indels: 0  
DB: 10 Gaps: 0  
US-08-836-455-2 (1-145) x US-09-924-099-20 (1-729)  
QY 17 GlyThrArgCysAspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGly 36  
Db 379 GGTGGCGGATCGGACATCCAGATGACCCAGTCTCCATCTCTCTATCTGCTCTCTGGGA 438  
QY 37 GlnArgValSerLeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeu 56  
Db 439 GAAGAGTCACTCTCACTTCTCGGGCAAGTCAGGACATTTAGTAAATATATCTGCTT 498  
QY 57 GlnGlnGluProAspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySer 76  
Db 499 CAACAGCAACAGATGGAACCTTTTAAACGCTGATCTACGCCACATCCAGTTTATGATCT 558  
QY 77 GlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSer 96  
Db 559 GGTGTCCCAAGAGGTTTCAGTGGCAGTAGGTCTGGTCCAGATTTATCTCTCACCATCAGC 618  
QY 97 SerLeuGluSerGluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyr 116  
Db 619 AGCTTCAGTCTGAAGATTTTGTAGACTATTACTGTCTACAATATGCTAGTCTCCGTAC 678  
QY 117 ThrPheGlyGlyThrLysLeuGluIleLys 127  
Db 679 ACGTTCGAGGGGGGACCAAGCTGGCAATAAAA 711



```
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 974
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(38)
; NAME/KEY: CDS
; LOCATION: (39)...(746)
; NAME/KEY: 3'UTR
; LOCATION: (750)...(974)
; NAME/KEY: sig_peptide
; LOCATION: (39)...(104)
US-09-859-053-29

Alignment Scores:
Pred. No.:      2,78e-50      Length:      974
Score:          473.00      Matches:      95
Percent Similarity: 76.92%      Conservative: 15
Best Local Similarity: 66.43%      Mismatches: 33
Query Match:      63.92%      Indels:      0
DB:              10      Gaps:      0

US-08-836-455-2 (1-145) x US-09-859-053-29 (1-974)

QY      1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuPheProGlyThrArgCys 20
Db      45 ATGAGGGTCCCGCTCAGCTCCTGGGGCTCCTGCTGCTGGTTCCAGGTTCCAGATGC 104
QY      21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40
Db      105 GACATCCAGATGACCCAGCTCCATCTCCGTGCTGTGCATCTGTAGGAGACAGAGTCACC 164
QY      41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGlnPro 60
Db      165 ATCACTGTGGCGCGAGTCAGGTAATAGCAGGTTGTTAGCCTGGTATCAGCAGAAACCA 224
QY      61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80
Db      225 GGGAAAGCCCCCTAAACTCCTGATCTATGTTGCATCCAGTTTGCAAAGTGGGGTCCCATCA 284
QY      81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100
Db      285 AGGTTTCAGCGCGAGTGGATCTGGCAGAGATTTCACTCTCACCATCAGCAGCCTGCAGCCT 344
QY      101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120
Db      345 GAAGATTTTGCACCTTACTATTGTCAACAGGCTACAGGTTCCCGTGGAGCGTTTCGGCCAA 404
QY      121 GlyThrLysLeuGluIleLysArgAlaAspAlaProThrValSerIlePheProPro 140
Db      405 GGGACCAAGGTGAATCAACAGCAACTGTGGCTGGCACCACATCTCTTCATCTTCCGCCCA 464
QY      141 SerSerLys 143
Db      465 TCTGATGAG 473
```

Search completed: August 31, 2003, 01:08:46  
Job time : 194.198 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 30, 2003, 22:11:21 ; Search time 1744.87 Seconds  
(without alignments)  
2019.725 Million cell updates/sec

Title: US-08-836-455-2  
Perfect score: 740  
Sequence: 1 MGAPQILGFLLLFPFTRC.....IKRADAAPTIVSIFPPSKLG 145

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US08836455/runat\_29082003\_132902\_22290/app\_query.fasta\_1.654  
-DB=EST -QPMF=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCFALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US08836455.ecgn\_1.1.3596.erunat\_29082003\_132902\_22290 -NCPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- EST:\*
- 1: em\_estba:\*
  - 2: em\_esthum:\*
  - 3: em\_estin:\*
  - 4: em\_estnu:\*
  - 5: em\_estov:\*
  - 6: em\_estpl:\*
  - 7: em\_estro:\*
  - 8: em\_htc:\*
  - 9: gb\_est1:\*
  - 10: gb\_est2:\*
  - 11: gb\_htc:\*
  - 12: gb\_est3:\*
  - 13: gb\_est4:\*
  - 14: gb\_est5:\*
  - 15: em\_estfun:\*
  - 16: em\_estom:\*
  - 17: em\_gss\_hum:\*
  - 18: em\_gss\_inv:\*
  - 19: em\_gss\_pln:\*
  - 20: em\_gss\_vrt:\*
  - 21: em\_gss\_fun:\*
  - 22: em\_gss\_mam:\*
  - 23: em\_gss\_mus:\*
  - 24: em\_gss\_pro:\*
  - 25: em\_gss\_rod:\*
  - 26: em\_gss\_phg:\*
  - 27: em\_gss\_vri:\*
  - 28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	630	85.1	630	10	BF138788	BF138788 601780387
2	618	83.5	926	12	BG962572	BG962572 602829953
3	546	73.0	714	12	BG963548	BG963548 602831226
4	526	71.1	374	13	BY083003	BY083003 BY083003
5	516	69.7	874	10	BG518527	BG518527 602578261
6	509	68.8	827	12	BI152061	BI152061 602916512
7	506	68.4	669	10	BF123422	BF123422 601760623
8	504	68.1	762	10	BF144806	BF144806 601791486
9	500	67.6	886	10	BG756818	BG756818 602710291
10	498.5	67.4	891	10	BF579422	BF579422 602093833
11	497	67.2	1459	13	BQ918407	BQ918407 AGENCOCURT
12	491	66.4	772	10	BF581992	BF581992 602099448
13	490	66.2	695	10	BE284224	BE284224 601099161
14	489	66.1	708	12	BI103114	BI103114 602889345
15	489	66.1	901	13	BQ943305	BQ943305 AGENCOCURT
16	489	66.1	961	10	BG341988	BG341988 602463136
17	486	65.7	725	10	BF580940	BF580940 602100636
18	484	65.4	788	14	CB957342	CB957342 AGENCOCURT
19	483	65.3	912	10	BF129120	BF129120 601811580
20	483	65.3	1100	10	BF663472	BF663472 602144635
21	482	65.1	464	2	HS070437	Bx480430 Homo sapi
22	482	65.1	685	10	BG518543	BG518543 602578277
23	481	65.0	767	14	CB985164	CB985164 AGENCOCURT
24	481	65.0	813	10	BG431143	BG431143 602498773
25	480	64.9	943	10	BF976230	BF976230 602245105
26	479	64.7	1038	10	BF757218	BF757218 602710591
27	478	64.6	992	10	BF976253	BF976253 602245130
28	477	64.5	651	14	CA946490	CA946490 n107b12.x
29	477	64.5	796	14	CB986053	CB986053 AGENCOCURT
30	477	64.5	830	12	BI455041	BI455041 603173343
31	477	64.5	977	14	CB984937	CB984937 AGENCOCURT
32	476	64.3	626	10	BF582283	BF582283 602101109
33	476	64.3	764	14	CB956251	CB956251 AGENCOCURT
34	476	64.3	807	14	CB957785	CB957785 AGENCOCURT
35	474	64.1	639	12	BM819912	BM819912 K-EST0088
36	474	64.1	724	12	BI837410	BI837410 603086702
37	473.5	64.0	903	13	BQ706785	BQ706785 AGENCOCURT
38	473	63.9	736	14	CB986552	CB986552 AGENCOCURT
39	473	63.9	745	14	CB958128	CB958128 AGENCOCURT
40	473	63.9	891	13	BX336281	BX336281 BX336281
41	472.5	63.9	767	14	CB958542	CB958542 AGENCOCURT
42	472	63.8	837	14	CB984807	CB984807 AGENCOCURT
43	471	63.6	740	14	CB956527	CB956527 AGENCOCURT
44	471	63.6	785	14	CB986236	CB986236 AGENCOCURT
45	471	63.6	797	14	CB987347	CB987347 AGENCOCURT

ALIGNMENTS

RESULT 1  
BF138788

LOCUS

DEFINITION 601780387F1 NCI\_CGAP\_Lu30 Mus musculus cdna clone IMAGE:4008404 5',  
630 bp mRNA linear EST 24-OCT-2000  
mRNA sequence.

ACCESSION BF138788

VERSION BF138788.1 GI:10977828

KEYWORDS EST.

SOURCE Mus musculus

ORGANISM Mus musculus (house mouse)

REFERENCE 1 (bases 1 to 630)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1AM9242 row: n column: 21  
 High quality sequence stop: 628.  
**FEATURES** Location/Qualifiers  
 source 1..630

source 1..630  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CZECH 1"  
 /db\_xref="taxon:10090"  
 /clone\_image="IMAGE:408404"  
 /tissue\_type="tumor, metastatic to mammary"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI-CGAP\_Lu30"  
 /note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"  
 BASE COUNT 162 a 158 c 150 g 160 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 4.83e-61 Length: 630  
 Score: 630.00 Matches: 122  
 Percent Similarity: 92.20% Conservative: 8  
 Best Local Similarity: 86.52% Mismatches: 11  
 Query Match: 85.14% Indels: 0  
 DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x BF138788 (1-630)

QY 3 AlapAaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCysAspIle 22  
 Db 3 GCTCCTGGCAGGTTTGGCTTCTGTGCTCTGTTCAGGTGCCAGATGACATC 62  
 QY 23 GlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSerLeuThr 42  
 Db 63 CAGATGACCAGTCTCCATCTCTATCTGCCTCTCTGGGAGAAAGATCAGTCTCACT 122  
 QY 43 CysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGluProAspGly 62  
 Db 123 TGGCGGCAAGTCAGGACATTTGGTGTAGCTTAACTGGTTTCAGCAGAAACCATGGA 182  
 QY 63 ThrIleLysArgLeuIleTyAlaThrSerSerLeuGlySerGlyValProLysArgPhe 82  
 Db 183 ACTATTAAACCCGATCTACGCCACATCCAGTTTAGATTCTGTGTGCCCAAAAGGTC 242  
 QY 83 SerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSerGluAsp 102  
 Db 243 AGTGGCAGTAGTCTGGGTCCAGATTATCTCTCACCATCAGCAGCCCTTGAGTCTGAAGAC 302  
 QY 103 PheValAlaTyTyCysLeuGlnTyAlaSerSerProTyrThrPheGlyGlyGlyThr 122  
 Db 303 TTTTCAGACATTACTCTTTTACAAATATGCTAGTTATCTCTCAAGTTTCGGTGTGGGACC 362  
 QY 123 LysLeuGlnIleLysArgAlaAspAlaAlaProThrValSerIlePheProProSerSer 142  
 Db 363 AAGCTGAGGCTGAAACGGGTGATGCTGCACCAACTGTATCCATCTTCCACCATTCAGT 422  
 QY 143 Lys 143  
 Db 423 GAG 425

**RESULT 2**  
 BG962572  
**LOCUS** 926 bp mRNA linear EST 12-JUN-2001  
**DEFINITION** 602829953F1 NCI-CGAP\_Co24 Mus musculus cDNA clone IMAGE:4984788 5', mRNA sequence.  
**ACCESSION** BG962572  
**VERSION** BG962572.1 GI:14350209  
**KEYWORDS** EST.  
**SOURCE** Mus musculus (house mouse)  
**ORGANISM** Mus musculus

**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NIH-MGC <http://mgc.nci.nih.gov/>.  
**AUTHORS** National Institutes of Health, Mammalian Gene Collection (MGC)  
**TITLE** Unpublished  
**JOURNAL** Contact: Robert Strausberg, Ph.D.  
**COMMENT** Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1AM10991 row: i column: 13  
 High quality sequence stop: 685.  
**FEATURES** Location/Qualifiers  
 source 1..926  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_image="IMAGE:4984788"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI-CGAP\_Co24"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI-CGAP Library."

BASE COUNT 230 a 259 c 205 g 232 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.79e-59 Length: 926  
 Score: 618.00 Matches: 117  
 Percent Similarity: 91.61% Conservative: 14  
 Best Local Similarity: 81.82% Mismatches: 12  
 Query Match: 83.51% Indels: 0  
 DB: 12 Gaps: 0

US-08-836-455-2 (1-145) x BG962572 (1-926)

QY 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
 Db 15 ATGAGGTTCTCTGCTCAGCTTTGGCTTCTGTGCTCTGCTTCAGGTACAGATGT 74  
 QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
 Db 75 GACATCCAGATGACCCAGTCTCCATCTCTCTCTCTCTGGGAGAAAGATCAGT 134  
 QY 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGlnGluPro 60  
 Db 135 CTCACCTTTGTCGGGCAAGTCAGGAAGTTGGTGTCTTCTTAAGTTGGCTTCAGCAAAAACCA 194  
 QY 61 AspGlyThrIleLysArgLeuIleTyAlaThrSerSerLeuGlySerGlyValProLys 80  
 Db 195 GATGGAACTTTTAAAGCCGCTGATCTACGCCACCACTTAGATTCTGGTGTCCCAAAA 254  
 QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
 Db 255 AGGTTTCAGTGGCAGTAGTCTGGGTCCAGATTATCTCTCACCCGTACAGCTTGTAGTCT 314

```

QY 101 GluAspPheValAlaTyrCysLeuGlnTyrAlaSerProTyrThrPheGlyGly 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 GAAGATTTCACACTATTAATCTCTACATATGCTATCGTGGAGTTCGGTGA 374

QY 121 GlyThrLysLeuGlnLeuLysArgAlaAspAlaProThrValSerIlePheProPro 140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 GGCACCGAGGTGGAAATCAACGGGTGATGCTGCACCAACTGTATCCATCTTCCACCA 434

QY 141 SerSerLys 143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435 TCCAGTGAG 443

RESULT 3
BG963548
LOCUS BG963548
DEFINITION 60283126F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4985791 5',
mRNA sequence.
ACCESSION BG963548
VERSION BG963548.1 GI:14351185
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NTH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10994 row: c column: 08
High quality sequence stop: 712.
FEATURES
    source
        1..714
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:4985791"
            /lab_host="DH10B (T1 phage-resistant)"
            /clone_lib="NCI_CGAP_Co24"
            /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.6 kb. Constructed by Life
            Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 189 a 178 c 167 g 180 t
ORIGIN
Alignment Scores:
Pred. No.: 8.55e-51 Length: 714
Score: 540.00 Matches: 108
Percent Similarity: 82.98% Conservative: 9
Best Local Similarity: 76.60% Mismatches: 24
Query Match: 12.97% Indels: 0
DB: 12 Gaps: 0

US-08-836-455-2 (1-145) x BG963548 (1-714)

QY 3 AlaProAlaGlnIleLeuGlyPheLeuLeuLeuLeuPheProGlyThrArgCysAspIle 22
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 22 GCTCCTGCTCAGTTCTTCTGGCATCTGTTGCTCTGTTCCAGGTGCAGATGTGACCTC 81

QY 23 GlnMetThrGlnSerProSerSerAlaSerLeuGlnArgValSerLeuThr 42
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 82 CAGATGATTCAGTCTCCATGCTCCATGTTTGCCTCTCTGGGAGACAGTGTCTCTCT 141

```

```

QY 43 CysArgAlaSerGlnAspIleGlyIleAsnLeuHistrPleuGlnGlnProAspGly 62
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142 TGTCGGACTAGTCAGGCATTAAGAGTAATTTAGACTGGTATCAGCAGAACACAGGTGA 201

QY 63 ThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLysArgPhe 82
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 202 ACTATTAAAGTCCTGATCTACTCCACATCAATTTAAATTTCTGTGTCTCCATCAAGGTT 261

QY 83 SerGlySerArgSerGlySerAspTyrSerThrIleSerSerLeuGlnSerGluAsp 102
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 AGTGGCAGTGGGTCTGGTCTCAGATTATCTCTCACCATCAGCAGCTTAGAGTCTGAAGAT 321

QY 103 PheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGlyThr 122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 TTTCGACACTATTACTGTCTACACGTCGTCATCGTGGAGTTCGGTGGCGGACCC 381

QY 123 LysLeuGlnIleLysArgAlaAspAlaProThrValSerIlePheProProSerSer 142
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 AAGCTGGAATCAACAGGGTGTGCTGCACCAACTGTATCCATCTTCCACCATCCAGT 441

QY 143 Lys 143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 GAG 444

RESULT 4
BG963003
LOCUS BG963003
DEFINITION etc. Mus musculus cDNA clone K630040M06 5', mRNA sequence.
ACCESSION BG963003
VERSION BG963003.1 GI:26193211
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Brad, D., Brusic, V., Chothia, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Kongaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltas, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perle, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, J. U., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sample, C. A., Setou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Venter, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sakai, D., Shibata,
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL 12466851
MEDLINE 22354683
PUBMED 12466851
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)

```

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@esc.riken.go.jp  
 URL: http://genome.esc.riken.go.jp/

Alzawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane  
 T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.,  
 Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,  
 Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,  
 M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct  
 Submission

Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.

## FEATURES

Location/Qualifiers

source

```
1..374
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone_lib="RIKEN full-length enriched, pooled tissues,
adult spleen, etc."
/note="(dev_stage=adult,tissue_type=spleen,sex=male),
(dev_stage=adult,tissue_type=kidney,sex=male),
(dev_stage=adult,tissue_type=testis,sex=male),
(dev_stage=adult,tissue_type=thymus,sex=male),
(dev_stage=adult,tissue_type=heart,sex=male),
(dev_stage=adult,tissue_type=colon,sex=male),
(dev_stage=adult,tissue_type=stomach,sex=male),
(dev_stage=adult,tissue_type=liver,sex=male),
(dev_stage=13 days embryo,tissue_type=whole body,sex=mix
),(dev_stage=14 days embryo,tissue_type=whole body,sex=mix
),(dev_stage=16 days embryo,tissue_type=whole body,sex=mix
),(dev_stage=17 days embryo,tissue_type=whole body,sex=mix
),(dev_stage=15 days pregnant,adult,tissue_type=amniot
,sex=female),(dev_stage=10 days neonate,tissue_type=brain
,sex=mix),(dev_stage=10 days neonate,tissue_type=thymus
,sex=mix),(dev_stage=10 days neonate,tissue_type=heart
,sex=mix)"
88 a 89 c 85 g 111 t 1 others
```

BASE COUNT

ORIGIN

Alignment Scores:

```
Pred. No.: 1..44e-49 Length: 374
Score: 526.00 Matches: 105
Percent Similarity: 94.74% Conservative: 3
Best Local Similarity: 92.11% Mismatches: 6
Query Match: 71.08% Indels: 0
DB: 13 Gaps: 0
```

US-08-836-455-2 (1-145) x BY083003 (1-374)

Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
 Db 33 ATGAGGGCTCTGCACAGATTTTGGCTCTCTGTGCTCTCTGTTCCAGGTACCAGATGT 92  
 Qy 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlnAlaSer 40  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 93 GACATCCAGATGACCCAGTCTCCATCCTCTCTATCTGCTCTCTGGGAGAAAGAGTCAGT 152  
 Qy 41 LeuThrCysArgAlaSerGlnAspIleGlytleAsnLeuHisTrpLeuGlnGlnGluPro 60  
 Db 153 CTCACCTTGTCCGGCAAGTCAGGACATTTGGTAGTAGTAACTGGCTTCAGCAGAACCA 212  
 Qy 61 AspGlyThrIleLysArgLeuIleTyAlaThrSerSerLeuGlySerGlyValProLys 80  
 Db 213 GATGGAACTATTAAACGCTGATCTACGCCACATCCAGTTTAGATTCTGGTGTNCCCAA 272  
 Qy 81 ArgPheSerGlySerArgSerGlySerAspTyTrSerLeuThrIleSerSerLeuGluSer 100  
 Db 273 AGGTTTCAGTGGCAGTAGTCTGGGTGAGATTATTTCTCCACATCAGCAGCCTTGAGTCT 332  
 Qy 101 GluAspPheValAlaTyTrCysLeuGlnTyAlaSerSer 114  
 Db 333 GAACATTTGTAGACTATTACTGTCTACAAATATGCTAGTCTCT 374

RESULT 5

BG518527

LOCUS

DEFINITION 602578261F1 NCI\_CGAP\_Lu29 Mus musculus cDNA clone IMAGE:3491798 5',  
 mRNA sequence.

ACCESSION

BG518527

VERSION

BG518527.1 GI:13513491

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LIAW8536 row: i column: 15

High quality sequence stop: 843.

Location/Qualifiers

1..874

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C2ECH II"

/db\_xref="taxon:10090"

/clone\_image="IMAGE:3491798"

/tissue\_type="spontaneous tumor, metastatic to mammary."

Stem cell origin."

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP\_Lu29"

/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"

BASE COUNT

ORIGIN

234 a 232 c 205 g 203 t

Alignment Scores:

Pred. No.: 5.73e-48 Length: 874

Score: 516.00 Matches: 102

Percent Similarity: 79.72% Conservative: 12

Best Local Similarity: 71.33% Mismatches: 29

Query Match: 69.73% Indels: 0

DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x BG518527 (1-874)



QY 1 MetClyAlaProLaGlnIleLeuGlyPheLeuLeuLeuLeuPheProGlyThrArgCys 20  
 DB 35 ATGAGGCCCTGCTGCTAGTTTGGGATCTTGTGCTCTGTTCCAGGTATCAGATGT 94  
 QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
 DB 95 GACATCAAGATGACCCAGTCTCATCTCCATGTCATGTCGCTGGGAGAGAGTCACT 154  
 QY 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuGlnHisTrpLeuGlnGlnPro 60  
 DB 155 ATCATTGCAAGCGGAGTCAAGACATTAAGCTATTAAAGCTGGTACACAGACCA 214  
 QY 61 AspGlyThrIleLysArgLeuIleTyAlaThrSerSerLeuGlySerGlyValProLys 80  
 DB 215 TGGAAATCTCTTAAGACCCCTGATCTATTATGCAACAGCTGGCAGATGCCCATCA 274  
 QY 81 ArgPheSerGlySerArgSerGlySerAspTySerLeuThrIleSerSerLeuGlySer 100  
 DB 275 AGATTACGTCAGTGGATCTGGCAGATATTCTTAACCATCAGCAGCCTGGAGTCT 334  
 QY 101 GluAspPheValAlaTyTrpCysLeuGlnTyAlaSerSerProTyTrpPheGlyGly 120  
 DB 335 GACGATACAGCAACTTATTACTCTACAGCATGGTGAGAGCCGTTACGTTCCGATCG 394  
 QY 121 GlyThrIleGlnIleLysArgAlaAspAlaProThrValSerIlePheProPro 140  
 DB 395 GGGACCAAGCTGGAATAAAACGGCTGATGCTGCACCAACTGTATCCATCTCCACCA 454  
 QY 141 SerSerLys 143  
 DB 455 TCCAGTGAG 463

RESULT 6  
 BIL52061 602916512F1 NCI\_CGAP\_Lu29 Mus musculus cDNA clone IMAGE:5066931 5',  
 LOCUS mRNA sequence.  
 DEFINITION BIL52061 GI:14612062  
 VERSION Mus musculus (house mouse)  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM1180 row: h column: 04  
 High quality sequence stop: 827.  
 Location/Qualifiers  
 1. 827  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="Czech II"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5066931"  
 /tissue\_type="spontaneous tumor, metastatic to mammary."  
 Stem cell origin.  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP\_Lu29"  
 /note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: Salt;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dr.  
 Library constructed by Life Technologies. Investigator

BASE COUNT 235 a 219 c 180 g 193 t  
 ORIGIN  
 Alignment Scores: 3.3e-47 Length: 827  
 Pred. No.: 509.00 Matches: 100  
 Score: 82.35% Conservative: 12  
 Best Local Similarity: 73.53% Mismatches: 24  
 Query Match: 68.78% Indels: 0  
 DB: 12 Gaps: 0  
 US-08-836-455-2 (1-145) x BIL52061 (1-827)  
 QY 8 LeuGlyPheLeuLeuLeuPheProGlyThrArgCysAspIleGlnMetThrGlnSer 27  
 DB 35 TTGGGTCTCTGTTGCTCTGTTTCAAGGTACCATGTCATATCCAGATGACACACT 94  
 QY 28 ProSerSerLeuSerAlaSerLeuGlyGlnArgValSerLeuThrCysArgAlaSerGln 47  
 DB 95 ACATCTCTCCCTGCTGCTCTCTGGGAGACAGAGTCAACCATCAGTTCAGGGCAAGTCAG 154  
 QY 48 AspIleGlyIleAsnLeuHisTrpLeuGlnGlnProAspGlyThrIleLysArgLeu 67  
 DB 155 GACATTAGCAATATTATTTAAACTGGTATCAGCAGAAACCATGGAAGTGTAAACTACTG 214  
 QY 68 IleTyAlaThrSerSerLeuGlySerGlyValProLysArgPheSerGlySerArgSer 87  
 DB 215 ATCTACTACATCAAGATTACACTCAGGAGTCCCATCAAGTTCAGTTCAGGCGGTCT 274  
 QY 88 GlySerAspTyTrpSerLeuThrIleSerSerLeuGluSerGluAspPheValAlaTyTrp 107  
 DB 275 GGGACAGATTATTCTCTCACTATTAGCAACTGGCAACAAGATATTGCCACTTACTTT 334  
 QY 108 CysLeuGlnTyAlaSerSerProTyTrpPheGlyGlyThrIleLysLeuGluIleLys 127  
 DB 335 TGGCAACAGGATGATAGCATCCGATATACGTCGGGACCAAGCTGGAATAAATAA 394  
 QY 128 ArgAlaAspAlaAlaProThrValSerIlePheProSerSerLys 143  
 DB 395 CGGCTGATGCTGCACCAACTGTATCCATCTTCCACCATCCAGTGAG 442  
 RESULT 7  
 BIL23422 601760623F1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:4023749 5',  
 LOCUS mRNA sequence.  
 DEFINITION BIL23422 GI:10962462  
 VERSION Mus musculus (house mouse)  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM9282 row: n column: 06  
 High quality sequence stop: 649.  
 Location/Qualifiers  
 1. 669  
 /organism="Mus musculus"  
 /mol\_type="mRNA"



[illegible]

RESULT 10	BF579422	891 bp	mrna	linear	EST 12-DEC-2000
LOCUS	BF579422	602093333F1	NCI_CGAP_Co24	Mus musculus	cdna clone IMAGE:4208144 5',
DEFINITION	mRNA sequence.				
ACCESSION	BF579422				
VERSION	BF579422.1				
KEYWORDS	EST.				
SOURCE	GI:11653134				
ORGANISM	Mus musculus (house mouse)				
	Mus musculus				

ORGANISM	REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
1 (bases 1 to 891)	
NIH-MGC <a href="http://mgc.ncbi.nlm.nih.gov/">http://mgc.ncbi.nlm.nih.gov/</a> .	
National Institutes of Health, Mammalian Gene Collection (MGC)	
Unpublished	
Contact: Robert Strausberg, Ph.D.	
Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a>	
Tissue procurement: Jeffrey F. Green M.D.	

Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAN9772 row: a column: 09  
 High quality sequence stop: 711.  
 Location/Qualifiers  
 1  
 891

FEATURES  
 source

```

source
i..891
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="420814"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site 2: SalI; Cloned unidirectionally; primer: Clono-qt

```

Site-2: SalI; cloned unidirectionally. Primer: Oligo dT. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library.

BASE COUNT	246 a	226 c	204 g	215 t
ORIGIN				
Alignment Scores:				
Pred. No.:	5.6e-46			891
Score:	498.50			101
Percent Similarity:	79.7%			13
Best Local Similarity:	70.6%			29
Query Match:	67.36%			1
DB:	10			0
				Gaps:
US-08-836-455-2 (1-145)	x	BF579422 (1-891)		

US-08-830-455-2 (1-145) X BF5/9422 (1-891)

1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuPheProGlyThrArgCys 20  
|||||  
13 ARGAGGACCCGCGCAGCTTCTTGGATCTTGTTGCTCTGGTTCCAGGTATCAAAATGT 72  
Db

21 AsprLleGlnMetThrGlnSerProSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
73 GACATCAAGATGACCAGCTCCATCTTCATGTCTGCATCTCTAGGAGAGAGTCACT 132

[illegible]

133 ATCACTTGC AAGCGAGTCAGGACATTAATAGCTATTTAAGCTGGTTCCAGCAGAAAACCA 192

DB6

09 AspGlyThrIleIevsArgIeuIleValAlaThrSerSerIeuClVserGlvAlproIvs 80

or aspartylmethyltransferase (EC 2.3.1.18) (10). The





CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM1122 row: i column: 11  
 High quality sequence stop: 706.  
 Location/Qualifiers  
 1. 708  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clones="IMAGE:5044690"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Kid14"  
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.75 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library. |"  
 BASE COUNT 198 a 182 c 155 g 173 t

## FEATURES

source

Alignment Scores:  
 Pred. No.: 4.96e-45 Length: 708  
 Score: 489.00 Matches: 96  
 Percent Similarity: 82.01% Conservative: 18  
 Best Local Similarity: 69.06% Mismatches: 25  
 Query Match: 66.08% Indels: 0  
 DB: 12 Gaps: 0

US-08-836-455-2 (1-145) x B1103114 (1-708)

Qy 5 AlaGlnLeuLeuGlyPheLeuLeuLeuPheProGlyThrArgCysAspIleGlnMet 24  
 Db 24 GTCAGTCTCTGGTCTCTGCTGCTGCTTTCAAGTACCAGATGATGTCACATG 83  
 Qy 25 ThrGlnSerProSerLeuSerAlaSerLeuGlyGlnArgValSerLeuThrCysArg 44  
 Db 84 ACACAGACTACATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 143  
 Qy 45 AlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeuGlnGluProAspGlyThrIle 64  
 Db 144 TCAAGTCAGAACATGTCCTACTTATTTGAAGTGGTTTCAGCAGAGCCAGATGGA 203  
 Qy 65 LysArgLeuIleTyrAlaThrSerSerLeuGlySerValProLysArgPheSerGly 84  
 Db 204 AAATCTCTATCTATTACATCAATCAAGATTACACTCAGAGTCCCATCAAGTTCA 263  
 Qy 85 SerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSerGluAspPheVal 104  
 Db 264 AGTGGGTCTGGACAAATTTTCTCTCACCATTAGCAACCTGGAAACAGAGATATTGG 323  
 Qy 105 AlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGlyThrLysLeu 124  
 Db 324 GCTATTATTTTGGCAACAGGTCTATACCTTCCTGGAGCGTTCGGTGGAGGCACCT 383  
 Qy 125 GluIleLysArgAlaAspAlaAlaProThrValSerIlePheProSerSerLys 143  
 Db 384 GAATCAACAGCGGTGATGCTGCACCACTGATCCATCTTCCACCATCCAGTGAG 440

RESULT 15  
 BQ943305  
 LOCUS  
 DEFINITION AGENCOURT 8880809 NCI\_CGAP\_Co24 Mus musculus CDNA clone  
 IMAGE:6397142 5', mRNA sequence.  
 BQ943305  
 ACCESSION BQ943305.1 GI:22358783  
 VERSION BQ943305.1  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## REFERENCE

AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 901)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: The Cepko Laboratory  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM1384 row: i column: 15  
 High quality sequence stop: 670.  
 Location/Qualifiers  
 1. 901  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clones="IMAGE:6397142"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Co24"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.6 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 231 a 243 c 199 g 227 t 1 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 6.74e-45 Length: 901  
 Score: 489.00 Matches: 97  
 Percent Similarity: 76.92% Conservative: 13  
 Best Local Similarity: 67.83% Mismatches: 33  
 Query Match: 66.08% Indels: 0  
 DB: 13 Gaps: 0

US-08-836-455-2 (1-145) x BQ943305 (1-901)

Qy 1 MetClyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
 Db 14 ATGAGGACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 73  
 Qy 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
 Db 74 GACATCAAGATGACCCAGTCTCCATCTCCATGTCATCTCTAGGAGCGGCTCACT 133  
 Qy 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGluPro 60  
 Db 134 ATCACTTGCAGGCGGAGTCAGGACATTTATTCCTTATTTAAATGTTTCCAGCAACCA 193  
 Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
 Db 194 GGGAGGCTCTCTAGACCCCTGATCATCAACACAGATTGATGGATGGGCTCCCTCA 253  
 Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
 Db 254 AGTTTCAGTGGCAGTGGATCTGGGCAAGATTTATCTCTCACCATCAACAGCTGGAGTGT 313  
 Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
 Db 314 GAAGATATGGAAATTTTATTTATTTGCTACAGTATGATGAGTTTCGCTCAGCTTCGGTGGT 373  
 Qy 121 GlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSerIlePhePro 140  
 Db 374 GGGACCAAGCTGGAGCTGAGACGGGCTGATGCTGCACCACTGATTCATCTTCCACCA 433  
 Qy 141 SerSerLys 143  
 Db 434 TCCAGTGAG 442

Search completed: August 31, 2003, 01:02:01  
Job time : 1748.87 secs

---

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 31, 2003, 00:02:06 ; Search time 2967.15 seconds  
(without alignments)  
1999.191 Million cell updates/sec

Title: US-08-836-455-2

Perfect score: 145

Sequence: 1 MGAQAQILGFLLLFPETRC.....IKRADAAPTVSIFPPSSKILG 145

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5773148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q/cgn2.1/USPTO.spool/US08836455/runat.29082003.132953.23045/app\_query.fasta.1.654  
-DB=GenEmbl -QFW=fastap -SUFFIX=p2noli.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US08836455.OCGN.1.1.3608.0/runat.29082003.132953.23045 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	145	100.0	435	6	AR164505 Sequence
2	145	100.0	435	6	BD085737 Methods o
3	125	86.2	381	10	AF124721 Mus muscu
4	33	22.8	387	6	AR169918 Sequence
5	31	21.4	535	6	BD021878 Humanized
6	30	20.7	438	6	E09035 cDNA encodi
7	30	20.7	438	6	I32991 Sequence 45
8	30	20.7	452	10	AB050077 Mus muscu
9	30	20.7	453	10	AB050084 Mus muscu
10	29	20.0	456	10	MUSIGKCNK
11	29	20.0	218	10	MMVJIG4
12	29	20.0	269	10	MMVJIG7
13	29	20.0	273	10	MMU21086
14	29	20.0	276	6	AR026090
15	29	20.0	276	6	AR026094
16	29	20.0	276	10	MMVJIG12
17	29	20.0	279	10	MMVIMRB11
18	29	20.0	283	10	AF144956
19	29	20.0	285	10	MMU29617
20	29	20.0	300	10	AF137626 Mus muscu
21	29	20.0	302	10	MMU55588
22	29	20.0	303	10	MUSIGLAFA
23	29	20.0	306	10	MUSIGKAF
24	29	20.0	321	10	AF163749
25	29	20.0	321	10	MUSIGKAA3
26	29	20.0	323	10	AY229938
27	29	20.0	324	6	AX722008
28	29	20.0	324	6	E54967 Peptide. 1/
29	29	20.0	324	6	I03643 Sequence 4
30	29	20.0	324	6	I07835 Sequence 4
31	29	20.0	324	10	MUSX
32	29	20.0	326	10	MMVJIG3
33	29	20.0	348	10	MMVJIG
34	29	20.0	354	10	AB089681 Mus muscu
35	29	20.0	381	10	AF045495
36	29	20.0	381	10	AF045508
37	29	20.0	381	10	AF045510
38	29	20.0	384	10	AB017434
39	29	20.0	390	10	MUSIKCC
40	29	20.0	407	6	E54981
41	29	20.0	443	10	MMU88675
42	29	20.0	447	6	AR083800
43	29	20.0	447	6	AR198719
44	29	20.0	447	6	AR275319
45	29	20.0	455	10	AB050082

#### ALIGNMENTS

RESULT 1

AR164505  
LOCUS AR164505 435 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 1 from patent US 6274143.  
ACCESSION AR164505  
VERSION AR164505.1 GI:162237555  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 435)  
AUTHORS Chatterjee, M. and Foon, K.A.  
TITLE Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11D10  
JOURNAL Patent: US 6274143-A 1 14-AUG-2001;  
FEATURES  
source Location/Qualifiers  
1..435  
/organism="unknown"  
BASE COUNT 100 a 111 c 102 g 122 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.91e-142 Length: 435  
Score: 145.00 Matches: 145  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-08-836-455-2 (1-145) x AR164505 (1-435)  
Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuLeuPheProGlyThrArgCys 20  
Dbb 1 ATGGGGCCCTGCTCAGATTCTTGGGTTCTTGTGCTTCTTGTTCAGGTACCATGT 60  
Qy 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
Dbb 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGACAAAGATCAGT 120  
Qy 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeuGlnGluPro 60  
Dbb 121 CTCACCTTGTCTGGGCAAGTCAGGACATTTGGTATTAACTTACATTGGCTTCAGCAGAACCA 180  
Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
Dbb 181 GATGGAACTATTAAACGCCCTGATCCGACATCCAGTTAGGTTCTGCTGACAAAGATCAGT 240  
Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
Dbb 241 AGGTTTCAGTGGCAGTAGTCTGGGTCCAGATTATTCTCTCACCATCAGCAGCTTGAGTCT 300  
Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaThrSerSerLeuGlySerGlyValProLys 120  
Dbb 301 GAAGATTTTGTAGCCTATTACTGTCTACAATATGCTAGTTCTCGGTACAGTTTCGGAGGG 360  
Qy 121 GlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSerIlePheProPro 140  
Dbb 361 GGGACCAAGCTGGAAATAAAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCCAACCA 420  
Qy 141 SerSerLysLeuGly 145  
Dbb 421 TCCAGTAAGCTTGGG 435  
RESULT 2  
BD085737  
LOCUS BD085737 435 bp DNA linear PAT 27-AUG-2002  
DEFINITION Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11D10.  
ACCESSION BD085737  
VERSION BD085737.1 GI:22631347  
KEYWORDS JP 2001523269-A/1.  
SOURCE unidentified  
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 435)  
AUTHORS Chatterjee, M. and Foon, K.A.  
TITLE Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11D10  
JOURNAL Patent: JP 2001523269-A 1 20-NOV-2001;  
COMMENT THE UNIVERSITY OF KENTUCKY RESEARCH FOUNDATION  
OS Unidentified  
PN JP 2001523269-A/1  
PD 20-NOV-2001  
PF 12-JUN-1998 JP 1999503252  
PR 13-JUN-1997 US 60/049540, 11-JUN-1998 US 09/096244 PI  
PC A61K39/395, A61K39/39//C07K16/42  
CC Strandedness: Single;  
CC Topology: Linear;  
CC Methods of delaying development of HMFG-associated tumors CC  
using  
CC anti-idiotypic antibody 11D10  
FH Key Location/Qualifiers  
FT CDS 1..435  
FT mat peptide 61.  
FEATURES  
source Location/Qualifiers  
1..435  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
BASE COUNT 100 a 111 c 102 g 122 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.91e-142 Length: 435  
Score: 145.00 Matches: 145  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-08-836-455-2 (1-145) x BD085737 (1-435)  
Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuLeuPheProGlyThrArgCys 20  
Dbb 1 ATGGGGCCCTGCTCAGATTCTTGGGTTCTTGTGCTTCTTGTTCAGGTACCATGT 60  
Qy 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
Dbb 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCTCTCTGGACAAAGATCAGT 120  
Qy 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisTrpLeuGlnGluPro 60  
Dbb 121 CTCACCTTGTCTGGGCAAGTCAGGACATTTGGTATTAACTTACATTGGCTTCAGCAGAACCA 180  
Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
Dbb 181 GATGGAACTATTAAACGCCCTGATCCGACATCCAGTTAGGTTCTGCTGACAAAGATCAGT 240  
Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
Dbb 241 AGGTTTCAGTGGCAGTAGTCTGGGTCCAGATTATTCTCTCACCATCAGCAGCTTGAGTCT 300  
Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaThrSerSerLeuGlySerGlyValProLys 120  
Dbb 301 GAAGATTTTGTAGCCTATTACTGTCTACAATATGCTAGTTCTCGGTACAGTTTCGGAGGG 360  
Qy 121 GlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSerIlePheProPro 140  
Dbb 361 GGGACCAAGCTGGAAATAAAACGGGCTGATGCTGCACCAACTGTATCCATCTTCCCAACCA 420  
Qy 141 SerSerLysLeuGly 145  
Dbb 421 TCCAGTAAGCTTGGG 435  
RESULT 3  
AF124721



```

C12N15/00,
PC C12N5/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
FEATURES
    source
        1..535
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
BASE COUNT 126 a 128 c 132 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 1.01e-22 Length: 535
Score: 31.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 21.38% Indels: 0
DB: 6 Gaps: 0
US-08-836-455-2 (1-145) x BD021878 (1-535)
QY 115 ProTyrThrPheGlyGlyThrLysLeuGlutLeuLysArgAlaAspAlaProThr 134
Db CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAACGGGCTGATGCTGCACCAACT 432
QY 135 ValSerIlePheProSerSerLysLeuGly 145
Db GTATCCATCTCCCAACCATCCAGTAAGCTTGGG 465
RESULT 6
E09035 438 bp RNA linear PAT 29-SEP-1997
LOCUS cdna encoding light chain variable region of mouse antiidiotype
DEFINITION antibody named Idiol7 against CLN-IgG idiotype.
ACCESSION E09035
VERSION E09035.1 GI:22025661
KEYWORDS JP 1995101999-A/7.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
    1 (bases 1 to 438)
Hagiwara.H. and Aozuka.Y.
AMINO ACID SEQUENCE OF ANTI-IDIOTYPE ANTIBODY TO ANTI-CANCER HUMAN
MONOCLONAL ANTIBODY AND DNA BASE SEQUENCE CODING FOR THE SAME
Patent: JP 1995101999-A 7 18-APR-1995;
JOURNAL HAGIWARA YOSHIHIDE
COMMENT OS Mus musculus (mouse)
PN JP 1995101999-A/7
PD 18-APR-1995
PF 06-OCT-1993 JP 1993272950
PI HAGIWARA HIDEAKI, AOZUKA YASUYUKI
PC C07K16/42,C12N5/10,C12N15/02,C12N15/09,C12P21/08,(C12P21/08,
PC C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source
    1..438
        /organism="Mus musculus"
        /tissue_type="spleen"
        /clone="17Kb1"
FT sig_peptide
    1..39
FT mat_peptide
    40..438
FT /note="Light chain variable region of mouse
FT antiidiotype
FT antibody named Idiol7 against CLN-IgG FT
FT idiotype"

```

```

FT misc_feature 40..105
FT /note="Fragment 1 of light chain variable FT
FT region of Idiol7"
FT misc_feature 106..153
FT /note="CDR1 of light chain variable region of
FT Idiol7"
FT misc_feature 154..198
FT /note="Fragment 2 of light chain variable FT
FT region of Idiol7"
FT misc_feature 199..219
FT /note="CDR2 of light chain variable region of
FT Idiol7"
FT misc_feature 220..315
FT /note="Fragment 3 of light chain variable FT
FT region of Idiol7"
FT misc_feature 316..339
FT /note="CDR3 of light chain variable region of
FT Idiol7"
FT misc_feature 340..369
FT /note="Fragment 4 of light chain variable FT
FT region of Idiol7"
FEATURES
    Location/Qualifiers
    source
        1..438
            /organism="Mus musculus"
            /mol_type="genomic RNA"
            /db_xref="taxon:10090"
BASE COUNT 104 a 120 c 112 g 102 t
ORIGIN
Alignment Scores:
Pred. No.: 9.31e-22 Length: 438
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.69% Indels: 0
DB: 6 Gaps: 0
US-08-836-455-2 (1-145) x E09035 (1-438)
QY 116 TyrThrPheGlyGlyThrLysLeuGlutLeuLysArgAlaAspAlaProThrVal 135
Db TACACGTTCCGAGGGGGACCAAGCTGGAATAAACGGGCTGATGCTGCACCAACTGTA 393
QY 136 SerIlePheProSerSerLysLeuGly 145
Db TCCATCTCCCAACCATCCAGTAAGCTTGGG 423
RESULT 7
I32991
LOCUS I32991 438 bp DNA linear PAT 06-FEB-1997
DEFINITION Sequence 45 from patent US 5589573.
ACCESSION I32991
VERSION I32991.1 GI:1823782
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 438)
AUTHORS Hagiwara.H. and Aotsuka.Y.
TITLE Amino acid sequences of anti-idiotypic antibodies against
anti-cancer human monoclonal antibody, and DNA base sequences
encoding those sequences
Patent: US 5589573-A 45 31-DEC-1996;
JOURNAL Location/Qualifiers
FEATURES
    source
        1..438
            /organism="unknown"
BASE COUNT 104 a 120 c 112 g 102 t
ORIGIN
Alignment Scores:
Pred. No.: 9.31e-22 Length: 438
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0

```

24

33

Qy 116 TyrThrPheGlyGlyGlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrVal 135

```

Db      363  TACAGTTCGGAGGGGACCAACAGCTGGAATAAACAAGGGCTGATGCTGCACCAACTGTA 422
Qy      136  SerIlePheProSerSerLysLeuGly 145
          |||||
Db      423  TCCATCTTCCACCATCCAGTAAGCTGGG 452

RESULT 10
MUSIGKCNK
LOCUS
DEFINITION Mouse Ig rearranged kappa-chain mRNA, clone AN12K.
ACCESSION M19914 J03832
VERSION M19914.1 GI:197035
KEYWORDS C-region; immunoglobulin light chain; immunoglobulin-kappa;
SOURCE processed gene.
ORGANISM Mus musculus (house mouse)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL 1 (bases 1 to 56)
REFERENCE Rule.G.S.
AUTHORS Unpublished (1988)
JOURNAL 2 (bases 1 to 456)
REFERENCE Leahy,D.J., Rule,G.S., Whittaker,M.M. and McConnell,H.M.
AUTHORS Sequences of 12 monoclonal anti-dinitrophenyl spin-label antibodies
TITLE for NMR studies
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (11), 3661-3665 (1988)
MEDLINE 88234486
PUBMED 3375235
COMMENT Original source text: Mus musculus (strain BALB/c, sub_species
domesticus) cDNA to mRNA.
Draft entry and computer-readable sequence [2] kindly submitted by
G.Rule, 20-JUL-1988.
FEATURES
source
1..456
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
/cell_type="mature B cell"
19..>456
/codon_start=1
/product="immunoglobulin kappa-chain"
/protein_id="AAA38889.1"
/db_xref="GI:466307"
/translat="GI:466307"
ITKASODVSTAVAWYQKPGQPKLLIYSASYRYGVDPDFTGSGSGTDFTTISV
QAEELAVYVYCHQYSSPYTFGGTKLEIKRAADAPTYSIFPPSS"
sig_peptide 19..90
mat_peptide 91..>456
BASE COUNT 118 a 114 c 108 g 116 t
ORIGIN
Alignment Scores:
Pred. No.: 9.66e-22 Length: 456
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20.69% Gaps: 0
10

US-08-836-455-2 (1-145) x MUSIGKCNK (1-456)
Qy 113 SerSerProThrPheGlyGlyThrLysLeuGluIleLysArgAlaAspAlaAla 132
|||||
Db 367 AGTTCTCCGTACAGTTCGAGGGGGGACCAAGCTGGAATAAACAAGGGCTGATGCTGCA 426

Qy 133 ProThrValSerIlePheProSerSer 142
|||||
Db 427 CCAACTGTATCATCTCTCCCAACCATCCAGT 456

RESULT 11

```

```

MMVJIG4
LOCUS
DEFINITION Mouse rearranged kappa immunoglobulin light chain (V,J).
ACCESSION X54757
VERSION X54757.1 GI:21998639
KEYWORDS Ig light chain; immunoglobulin.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL 1 (bases 1 to 218)
REFERENCE Harada,K. and Yamagishi,H.
AUTHORS Lack of feedback inhibition of V kappa gene rearrangement by
TITLE productively rearranged alleles
JOURNAL J. Exp. Med. 173 (2), 409-415 (1991)
MEDLINE 91108338
PUBMED 1988542
REFERENCE 2 (bases 1 to 218)
AUTHORS Yamagishi,H.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-1990) Yamagishi H., Dept of Biophysics Faculty of
Science Kyoto University, Sakyo-ku Kyoto 606, Japan
FEATURES
source
1..218
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="BALB/c-nu/nu"
/db_xref="taxon:10090"
/cclone="MSI-N102"
/cell_type="splenocyte"
/tissue_type="spleen"
/cclone_lib="lambda gt11"
/dev_stage="8 week old"
<1..>218
/notes="Ig light chain"
/codon_start=3
/protein_id="CAA38556.1"
/db_xref="GI:21998640"
/translat="GI:21998640"
EDFVDVYVYQYASSPFTFGSGTKLEIKR"
mat_peptide <1..>218
prim_transcript <1..>218
misc_feature 1..178
misc_feature 179..216
BASE COUNT 59 a 49 c 48 g 62 t
ORIGIN
Alignment Scores:
Pred. No.: 5.46e-21 Length: 218
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x MMVJIG4 (1-218)
Qy 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95
|||||
Db 60 TCTGGTGTCCCAAAAGGTTTCAGTGGCAGTAGGTCTGGTGCAGATTATTCTCTCACCATC 119

Qy 96 SerSerLeuGluSerGluAspPheVal 104
|||||
Db 120 AGCAGCCTTCAGTCTGAAGATTGTGA 146

RESULT 12
MMVJIG7
LOCUS
DEFINITION Mouse rearranged kappa immunoglobulin light chain (V,J).

```



```
source 1..276
/organism="unknown"
BASE COUNT 68 a 65 c 63 g 80 t
ORIGIN

Alignment Scores:
Pred. No.: 6.77e-21 Length: 276
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 6 Gaps: 0

US-08-836-455-2 (1-145) x AR026090 (1-276)
QY 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95
Db 142 TCTGGTGTCCCAAAAGGTTTCAGTGGCAGTAGGCTGGGTCTGAGATTATCTCTCACCATC 201
QY 96 SerSerLeuGluSerGluAspPheVal 104
Db 202 AGCAGCCTTGAGTCTGAAGATTGTA 228

RESULT 15
AR026094 LOCUS 276 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 52 from patent US 5855885.
ACCESSION AR026094
VERSION AR026094.1 GI:5936934
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 276)
AUTHORS Smith,R., McCafferty,J., Chiswell,D., Darsley,M.J., Fitzgerald,K.,
TITLE Isolation and production of catalytic antibodies using phage
JOURNAL technology
FEATURES Patent: US 5855885-A 52 05-JAN-1999;
Location/Qualifiers
source 1..276
/organism="unknown"
BASE COUNT 67 a 65 c 64 g 80 t
ORIGIN

Alignment Scores:
Pred. No.: 6.77e-21 Length: 276
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 6 Gaps: 0

US-08-836-455-2 (1-145) x AR026094 (1-276)
QY 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95
Db 142 TCTGGTGTCCCAAAAGGTTTCAGTGGCAGTAGGCTGGGTCTGAGATTATCTCTCACCATC 201
QY 96 SerSerLeuGluSerGluAspPheVal 104
Db 202 AGCAGCCTTGAGTCTGAAGATTGTA 228

Search completed: August 31, 2003, 03:58:12
Job time : 2968.15 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 30, 2003, 22:12:36 ; Search time 217.987 Seconds  
(without alignments)  
1795.608 Million cell updates/sec

Title: US-08-836-455-2

Perfect score: 145

Sequence: 1 MGAQAQILGFLLLFPGR.....IKRAAPATVSIPIPPSSKLG 145

Scoring table:

OLIGO.  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5103490

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh  
-Q=/cnp2\_1/USPTO.spool/US08836455/runat\_29082003\_132953\_23037/app\_query.fasta.1.654  
-DB\_N\_Geneseq\_19Jun03 -QFMT=fastep -SUFFIX=p2noli.rng -MINMATCH=0.1 -LOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=OLIGO -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US08836455\_QCGN\_1\_1\_401\_@runat\_29082003\_132953\_23037 -NCPU=6 -ICPU=3  
-NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_19Jun03:\*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*
- 25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	145	100.0	435	18	AAT85149	Murine monoclonal
2	145	100.0	435	20	AAV83772	Antibody 11D10 lig
3	145	100.0	435	25	AA51273	Murine 11D10 antibo
4	33	22.8	387	18	AAT77851	Murine anti-human
5	33	22.8	387	24	AAD32138	Murine 44H104 mab
6	31	21.4	535	19	AAV20086	Consensus DNA sequ
7	30	20.7	438	16	AAQ90431	DNA encoding anti-
8	29	20.0	276	20	AAQ00879	Mouse derived RT3
9	29	20.0	276	20	AAQ00875	Mouse derived RT3
10	29	20.0	321	21	AAQ38908	520C9 hybridoma VL
11	29	20.0	324	21	AAQ49534	Mouse anti-IL-18 a
12	29	20.0	390	25	AAD52607	Escherichia coli 1
13	29	20.0	407	21	AAQ49548	Mouse light chain
14	29	20.0	447	17	AAT31332	Anti-idiotypic mono
15	29	20.0	447	20	AAT31365	MAB 1A7 light chai
16	29	20.0	447	20	AAQ89552	Light chain variab
17	29	20.0	450	4	AAQ30165	Sequence encoding
18	29	20.0	711	21	AAQ49542	pESCFV#125-2H reco
19	29	20.0	729	21	AAQ49543	pESCFV#125-2H.HT r
20	29	20.0	739	14	AAQ46084	Sequence encoding
21	29	20.0	739	17	AAT36880	520C9 anti-c-erbB-
22	29	20.0	739	19	AAV21798	520C9 anti-c-erbB-
23	29	20.0	739	20	AAV63399	520C9 sfv DNA sequ
24	29	20.0	756	24	AAQ97136	P4-3 single chain
25	29	20.0	771	24	AAQ97142	P5-10 single chain
26	29	20.0	1497	24	AAQ97145	3B10xP4-3 bispecif
27	29	20.0	1605	14	AAQ46086	Sequence encoding
28	28	19.3	321	21	AAQ38909	650E2 hybridoma VL
29	28	19.3	456	22	AAQ66996	Filamentous phage
30	28	19.3	639	10	AAQ91657	Chimeric antibody
31	28	19.3	642	18	AAQ85091	Mouse monoclonal a
32	28	19.3	651	21	AAQ44346	Human secreted exp
33	28	19.3	652	17	AAT87818	Antibody 3c2 light
34	28	19.3	654	25	ACC44908	TSH receptor antib
35	28	19.3	654	25	ACC44909	TSH receptor antib
36	28	19.3	678	21	AAQ27849	WOW-1 Fab light ch
37	28	19.3	698	25	ABX16574	Mouse DNA encoding
38	28	19.3	723	16	AAQ92503	Mouse antibody F4-
39	28	19.3	738	21	AAQ61037	Nucleotide sequenc
40	28	19.3	867	22	AAQ67002	Filamentous phage
41	28	19.3	979	14	ABQ52791	Murine m166 antibo
42	27	18.6	366	13	AAQ27140	ICAM-1 inhibiting
43	27	18.6	429	18	AAT70811	Mouse anti-idiotyp
44	27	18.6	465	16	AAQ85387	MAB 4197X light ch
45	27	18.6	882	14	AAQ48038	Monoclonal antibod

ALIGNMENTS

RESULT 1

AAT85149  
ID AAT85149 standard; cDNA; 435 BP.

XX AAT85149;

DT 25-MAR-2003 (updated)

DT 04-JAN-1998 (first entry)

XX Murine monoclonal anti-idiotypic antibody 11D10 VL cDNA.

XX Monoclonal antibody 11D10; anti-idiotypic antibody; mucin;

KW human milk fat globule; HMFG; tumour; breast cancer; vaccine; ss.

XX Mus musculus.

XX Key

XX Location/Qualifiers

FT sig\_peptide 1..60  
 FT /\*tag= a  
 FT mat\_peptide 61..435  
 FT /\*tag= b  
 XX  
 PN WO9722699-A2.  
 XX  
 PD 26-JUN-1997.  
 XX  
 XX 19-DEC-1996; 96WO-US20757.  
 XX  
 XX 20-DEC-1995; 95US-0575762.  
 PR 26-JAN-1996; 96US-0591965.  
 PR 13-DEC-1996; 96US-0766350.  
 XX  
 XX (KENT ) UNIV KENTUCKY.  
 XX  
 XX Chatterjee M, Chatterjee SK, Foon KA;  
 XX  
 DR WPI; 1997-341690/31.  
 DR P-PSDB; AAW27119.  
 XX  
 XX Monoclonal anti-idiotypic antibody 11D10 - elicits immune response  
 PT against human milk fat globule disease associated tumours,  
 PT especially breast cancer  
 XX  
 XX Claim 11; Page 94; 130pp; English.  
 XX  
 XX This cDNA sequence encodes the light chain variable region VL  
 CC (AAW85149) of monoclonal anti-idiotypic antibody 11D10 produced by  
 CC hybridoma cell line ATCC 12020. 11D10 was obtained by immunising  
 CC naive mice with MC-10 anti-HMGF antibody to obtain an anti-idiotypic  
 CC response. It elicits an immune response against a specific epitope  
 CC of a high mol.wt. mucin of human milk fat globule (HMFG). It  
 CC induces an immunological response to HMGF in mice, rabbits, monkeys  
 CC and patients with advanced HMGF-associated tumours. Pharmaceutical  
 CC compositions and vaccines comprising 11D10, 11D10 polypeptides  
 CC and/or 11D10 polynucleotides are claimed. Also claimed are  
 CC diagnostic kits and methods of using 11D10, 11D10 polypeptides  
 CC and/or 11D10 polynucleotides, including methods of treating HMGF-  
 CC associated tumours.  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 XX  
 XX Sequence 435 BP; 100 A; 111 C; 102 G; 122 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. NO.: 1.09e-124 Length: 435  
 Score: 145.00 Matches: 145  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 18 Gaps: 0

US-08-836-455-2 (1-145) x AAT85149 (1-435)

Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
 Db 1 ATGGGGCCCCCTGCTCAGATTCTTGGGTTCTTGTGCTTCTGTTCCAGGTACCATGATG 60  
 Qy 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlnArgValSer 40  
 Db 61 GACATCCAGATGACCCAGTCCATCTCTTATCTGCTCTCTGCGACAAAGAGTCACT 120  
 Qy 41 LeuThrCysArgAlaSerGlnAspIleGlyLeuAsnLeuHisThrLeuGlnGluPro 60  
 Db 121 CTCACCTTGTGGGCAAGTCAGGACATTGGTATTAACTTACATTGGCTTCAGCAGAACCA 180  
 Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerSerLeuGlySerGlyValProLys 80  
 Db 181 GATGGAACCTATTAAACCCCTGATCTACGCCATCCAGTTAGGTTCTGTTGCCCAAA 240  
 Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100

Db 241 AGGTTTCAGTGGCAGTAGGTCTGGGTCTCAGATTATTTCTCTCACCATCAGCAGCCTTGAGTCT 300  
 Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerProTyrThrPheGlyGly 120  
 Db 301 GAAGATTTTCTAGCCTATTACTGCTACATATGCTAGTCTCGTACACGTTCCGGAGG 360  
 Qy 121 GlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSerIlePheProPro 140  
 Db 361 GGGACCAAGCTGGAATAAAGCGGTGATGCTGCACCAACTGTATCCATCTTCCCAACA 420  
 Qy 141 SerSerLysLeuGly 145  
 Db 421 TCCAGTAAGCTTGGG 435  
 RESULT 2  
 AAV83772  
 ID AAV83772 standard; cDNA; 435 BP.  
 XX  
 AC AAV83772;  
 XX  
 DT 16-MAR-1999 (first entry)  
 XX  
 DE Antibody 11D10 light chain variable region coding sequence.  
 XX  
 XX Murine; mouse; antibody; light chain; variable region; anti-idiotypic; ss;  
 KW human milk fat globule; tumour; ovary; lung; pancreas; carcinoma; breast.  
 XX  
 XX Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..435 /\*tag= a  
 FT /product= "antibody 11D10 light chain variable region"  
 FT /transl\_except= (pos:163..165, aa:Thr)  
 FT /note= "no stop codon is given at the 3' end of the  
 FT sequence"  
 FT sig\_peptide 1..60  
 FT /\*tag= b  
 FT mat\_peptide 61..435  
 FT /\*tag= c  
 XX  
 XX WO9856419-A1.  
 XX  
 PD 17-DEC-1998.  
 XX  
 XX 12-JUN-1998; 98WO-US12250.  
 XX  
 PR 11-JUN-1998; 98US-0096244.  
 PR 13-JUN-1997; 97US-0049540.  
 XX  
 PA (KENT ) UNIV KENTUCKY RES FOUND.  
 XX  
 XX Chatterjee M, Foon KA;  
 XX  
 DR WPI; 1999-060029/05.  
 DR P-PSDB; AAW87593.  
 XX  
 PT Delaying development of, or treating, HMGF-associated tumours -  
 PT using anti-idiotypic antibody 11D10 raised against antibodies to  
 PT human milk fat globule protein  
 XX  
 PS Disclosure; Fig 1; 54pp; English.  
 XX  
 CC This sequence represents the coding sequence for the murine antibody  
 CC 11D10 light chain variable region. This anti-idiotypic antibody is used  
 CC to delay the development of, or treat, a human milk fat globule (HMFG)  
 CC associated tumour in an individual having low tumour burden.  
 CC The antibody 11D10 is used to prevent the recurrence of HMGF-associated  
 CC tumours e.g. ovarian, non-small cell lung and pancreatic carcinoma,  
 CC especially for treating breast tumours.  
 XX  
 SQ Sequence 435 BP; 100 A; 111 C; 102 G; 122 T; 0 other;

## Alignment Scores:

Pred. No.: 1,09e-124 Length: 435  
 Score: 145.00 Matches: 145  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

US-08-836-455-2 (1-145) x AAV83772 (1-435)

QY 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
 DB 1 ATGGGGCCCTCCTCAGATCTTGGGTTCTTGTGCTCTGTTTCCAGGTACAGATGT 60  
 QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
 DB 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCCTCTCTGGACAAAGAGTCAGT 120  
 QY 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGluPro 60  
 DB 121 CTCACCTGTGGCAAGTCAGGACATTTGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180  
 QY 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerLeuGlySerGlyValProLys 80  
 DB 181 GATGGAACTATTAAACGCCGTGATCTAGCCACATCCAGTTTGTGTTGTCGCCAAA 240  
 QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
 DB 241 AGGTTTCAGTGGCAGTAGTCTGGTCTGAGATTTATCTCACCATCAGCAGCCTTGAGTCT 300  
 QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
 DB 301 GAAGATTTGTAGCCCTATTACTCTCACAATATGCTAGTTCCTCGTACAGTTTCGGAGGG 360  
 QY 121 GlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSerIlePheProPro 140  
 DB 361 GGCACCAAGCTGGAATAAAACGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
 QY 141 SerSerLysLeuGly 145  
 DB 421 TCCAGTAAGCTTGGG 435

## RESULT 3

ID AAL51273 standard; cDNA; 435 BP.

XX AC AAL51273;

XX DT 20-MAR-2003 (first entry)

XX DE Mouse 11D10 antibody light chain variable region coding sequence.

XX KW Mouse; murine; gene; ss; vaccine; tumour; human milk fat globules; HMFG;

XX KW carcinoembryonic antigen; CEA; 11D10; 3H1; HMFG-associated tumour;

XX KW CEA-associated tumour; anti-idiotype antibody.

XX OS Mus musculus.

XX FH Key Location/Qualifiers

FT CDS 1..435

FT /\*tag= a

FT /partial

FT /product= "Mouse 11D10 anti-idiotype antibody light chain

FT variable region"

FT /note= "No stop codon is given"

FT sig\_peptide 1..60

FT /\*tag= b

FT mat\_peptide 61..435

FT /\*tag= c

XX WO200292012-A2.

XX PN 21-NOV-2002.

XX PD

XX

XX PF 17-MAY-2002; 2002WO-US15840.

XX PR 17-MAY-2001; 2001US-0861294.

XX PA (KENT ) UNIV KENTUCKY RES FOUND.

XX PI Chatterjee M, Foon KA;

XX PS WPI: 2003-129216/12.

XX DR P-PSDB; AAO16292.

XX PT Use of anti-idiotypic antibodies for human milk fat globules (HMFG) - or

XX PT carcinoembryonic antigen (CEA)-associated tumor for delaying the

XX PT development of, or treating a HMFG- or CEA-associated tumor (e.g.

XX PT breast tumor) in humans

XX PS Disclosure; Fig 1; 98pp; English.

XX CC The invention comprises a method for delaying the development of, or

XX CC treating a tumour that is associated with human milk fat globules (HMFG)

XX CC or carcinoembryonic antigen (CEA). The method of the invention involves

XX CC administering an anti-idiotypic antibody for HMFG (e.g. 11D10), and an

XX CC anti-idiotypic antibody for CEA (e.g. 3H1). The method is useful for

XX CC delaying the development, of or treating HMFG/CEA-associated tumours. The

XX CC present cDNA sequence encodes the light chain variable region of the

XX CC mouse 11D10 anti-idiotype antibody.

XX SQ Sequence 435 BP; 100 A; 111 C; 102 G; 122 T; 0 other;

## Alignment Scores:

Pred. No.: 1,09e-124 Length: 435  
 Score: 145.00 Matches: 145  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 25 Gaps: 0

US-08-836-455-2 (1-145) x AAL51273 (1-435)

QY 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20  
 DB 1 ATGGGGCCCTCCTCAGATCTTGGGTTCTTGTGCTCTGTTTCCAGGTACAGATGT 60  
 QY 21 AspIleGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40  
 DB 61 GACATCCAGATGACCCAGTCTCCATCTCTTATCTGCCTCTCTGGACAAAGAGTCAGT 120  
 QY 41 LeuThrCysArgAlaSerGlnAspIleGlyIleAsnLeuHisTrpLeuGlnGluPro 60  
 DB 121 CTCACCTGTGGCAAGTCAGGACATTTGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180  
 QY 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerLeuGlySerGlyValProLys 80  
 DB 181 GATGGAACTATTAAACGCCGTGATCTAGCCACATCCAGTTTGTGTTGTCGCCAAA 240  
 QY 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100  
 DB 241 AGGTTTCAGTGGCAGTAGTCTGGTCTGAGATTTATCTCACCATCAGCAGCCTTGAGTCT 300  
 QY 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerSerProTyrThrPheGlyGly 120  
 DB 301 GAAGATTTGTAGCCCTATTACTCTCACAATATGCTAGTTCCTCGTACAGTTTCGGAGGG 360  
 QY 121 GlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSerIlePheProPro 140  
 DB 361 GGCACCAAGCTGGAATAAAACGGCTGATGCTGCACCAACTGTATCCATCTTCCACCA 420  
 QY 141 SerSerLysLeuGly 145  
 DB 421 TCCAGTAAGCTTGGG 435

RESULT 4

```
AAAT77851
ID  AAAT77851 standard; cDNA; 387 BP.
AC  AAAT77851;
XX  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DT  03-NOV-1997 (first entry)
DE  Murine anti-human class II monoclonal antibody 44H104 VL chain cDNA.
KW  Antibody; light chain; variable region; hybridoma cell line 44H104;
KW  immune response; enhance; stimulate; vaccine; immunodiagnosis;
KW  antigen delivery; ss.
XX  Mus musculus.
XX  Key Location/Qualifiers
XX  CDS 1..387
XX  FT /*tag= a
XX  FT /note= "Encodes 44H104 light chain variable region,
XX  FT including secretion signal; termination
XX  FT codon not given"
XX  PN WO9640941-A1.
XX  PD 19-DEC-1996.
XX  PF 07-JUN-1996; 96WO-CA00400.
XX  PR 07-JUN-1995; 95US-0483576.
XX  PA (CONN-) CONNAUGHT LAB LTD.
XX  PI Anand NN, Barber BH, Caterini JE, Cates GC, Klein MH;
XX  WPI: 1997-077271/07.
XX  P-PSDB; AAW22537.
XX  Recombinant conjugate antibody mol., modified for delivering an
XX  antigen - elicits enhanced immune response without the use of
XX  adjuvant to generate antibodies which are useful in vaccines or
XX  immuno:diagnosis
XX  Example 1; Fig 1A; 64pp; English.
XX  Novel recombinant conjugate antibody molecules comprise a monoclonal
XX  antibody specific for a surface structure of antigen presenting
XX  cells (APC), genetically modified to contain at least one antigen
XX  exclusively at one or more preselected sites. The conjugate is capable
XX  of delivering the antigen to APC and eliciting an immune response to
XX  the antigen. The new conjugates are useful as vaccines and are able
XX  to elicit an enhanced immune response without the use of an adjuvant.
XX  In a specific example, a conjugate was constructed using the murine
XX  anti-human class II monoclonal antibody secreted by hybridoma
XX  44H104. The peptide CLTB36 was chosen as antigen; it consists of
XX  a tandemly linked T and B cell epitope derived from HIV MN strain.
XX  The present sequence encodes the light chain variable region which
XX  was PCR amplified from 44H104 and used in the preparation of a
XX  conjugate with antigen CLTB36.
XX  Sequence 387 BP; 90 A; 95 C; 90 G; 112 T; 0 other;
XX  Alignment Scores:
XX  Pred. No.: 2.25e-21 Length: 387
XX  Score: 33.00 Matches: 33
XX  Percent Similarity: 100.00% Conservative: 0
XX  Best Local Similarity: 100.00% Mismatches: 0
XX  Query Match: 22.76% Indels: 0
XX  DB: 18 Gaps: 0
XX  US-08-836-455-2 (1-145) x AAAT77851 (1-387)
XX  QY 15 PheProGlyThrArgCysAspIleGlnMetThrGlnSerProSerLeuSerAlaSer 34
XX  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 49 TTTCAGGTACCATGATGTGACATCCAGATGATCCAGATGATCCAGTCTCATCTCTTATCTGCCTCT 108
QY 35 LeuGlyGlnArgValSerLeuThrCysArgAlaSerGln 47
XX  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 109 CTGGACAAAGAGTCAGTCTCATTGTCGGGCAAGTCAG 147
RESULT 5
AAD32138
ID  AAD32138 standard; DNA; 387 BP.
XX  AAD32138;
XX  AAD32138;
XX  18-JUN-2002 (first entry)
XX  Murine 44H104 mab variable light chain (VL) DNA.
XX  Murine; mab; light chain; VL; conjugate antibody; antigen delivery;
XX  immune system; vaccine; detecting agent; antibacterial; gene; ds.
XX  Mus sp.
XX  Key Location/Qualifiers
XX  CDS 1..387
XX  FT /*tag= a
XX  FT /product= "Murine 44H104 mab VL"
XX  FT /transl_except= (pos:277..279, aa:Thr)
XX  FT /note= "This translation exception occurs while
XX  FT decoding for murine mab VL alternative version
XX  FT (AAE20204); CDS does not include stop codon"
XX  FT /partial
XX  PN US2002025315-A1.
XX  PD 28-FEB-2002.
XX  PF 14-JAN-1998; 98US-0007093.
XX  PR 14-JAN-1998; 98US-0007093.
XX  PA (ANAN/) ANAND N N.
XX  PA (BARB/) BARBER B H.
XX  PA (CATE/) CATES G A.
XX  PA (CATE/) CATERINI J E.
XX  PA (KLEI/) KLEIN M H.
XX  PI Anand NN, Barber BH, Cates GA, Caterini JE, Klein MH;
XX  WPI: 2002-267519/31.
XX  P-PSDB; AAE20200, AAE20204.
XX  Novel recombinant conjugate antibody, useful as a vaccine against
XX  pathogens having a specific antigen, comprises a monoclonal antibody
XX  specific for an antigen presenting cell surface structure -
XX  Example 1; Fig 1A; 28pp; English.
XX  The invention relates to a recombinant conjugate antibody, comprising
XX  a monoclonal antibody specific for a surface structure of antigen
XX  presenting cells genetically modified to contain an antigen moiety
XX  for the purpose of delivery of the antigen moiety to antigen-
XX  presenting cells of the immune system. The conjugate antibody is
XX  formulated as a vaccine to protect a host against a disease caused
XX  by a pathogen expressing the antigen. The antibody is useful as a
XX  detecting agent. The present sequence is murine 44H104 mab variable
XX  light chain (VL) DNA.
XX  Sequence 387 BP; 90 A; 95 C; 90 G; 112 T; 0 other;
XX  Alignment Scores:
XX  Pred. No.: 2.25e-21 Length: 387
XX  Score: 33.00 Matches: 33
XX  Percent Similarity: 100.00% Conservative: 0
XX  Best Local Similarity: 100.00% Mismatches: 0
```

```

Query Match:      22.76%      Indels:      0
DB:               24          Gaps:         0

US-08-836-455-2 (1-145) x AAD32138 (1-387)

QY 15 PheProGlyThrArgCysAspIleGlnMetThrGlnSerProSerSerLeuSerAlaSer 34
DB 49 TTTCCAGGTACCATGTCATCCAGATCCAGATCCAGTCTCCATCTCTTATCTGCTCT 108

QY 35 LeuGlyGlnArgValSerLeuThrCysArgAlaSerGln 47
DB 109 CTGGGACAAAGATCAGTCTCACTTGTGGGCAAGTCAG 147

RESULT 6
AAV20086
ID AAV20086 standard; DNA; 535 BP.
XX
AC AAV20086;
XX
XX 14-JUL-1998 (first entry)
XX
DE Consensus DNA sequence of the murine variable light chain region.
XX
KW Mouse; Act-1 antibody; human alpha4-beta7 integrin;
KW Muscosal adressin cell adhesion molecule-1; MadCAM-1;
KW humanised antibody; murine antigen binding region; inhibition;
KW leukocyte infiltration of tissue; treatment; inflammatory disease;
KW inflammatory bowel disease; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 16..435
FT FT /*tag= a
FT FT /note= "no stop codon given"
FT FT
FT sig_peptide 16..75
FT FT /*tag= b
FT mat_peptide 76..435
FT FT /*tag= c
XX
XX WO9806248-A2.
XX
XX 19-FEB-1998.
XX
XX 06-AUG-1997; 97WO-US13884.
XX
XX 15-AUG-1996; 96US-0700737.
XX
XX (LEUK-) LEUKOSITE INC.
XX
XX Bendig MM, Jones ST, Newman W, Ponath PD, Ringler DJ;
XX Saldanha J;
XX
XX WPI; 1998-159172/14.
XX P-PSDB; AAW53817.
XX
XX Humanised immunoglobulin reactive with alpha-4-beta-7 integrin -
XX used for treating inflammatory disease, pancreatitis, diabetes,
XX asthma, graft versus host disease and sarcoidosis
XX
XX Example 1; Fig 3; 145pp; English.
XX
XX The present sequence represents the consensus nucleotide sequence
XX comprising the variable region of murine Act-1 antibody determined from
XX several independent mouse light chain variable region clones. Act-1 is
XX active against human alpha4-beta7 integrin. Muscosal adressin cell
XX adhesion molecule-1 (MadCAM-1) is a ligand of this particular integrin.
XX The Act-1 antibody interferes with alpha4-beta7 integrin binding to
XX MadCAM-1, which is present of high endothelial venules in mucosal
XX lymph nodes. Variable regions were amplified from DNA encoding Act-1
XX using degenerate PCR primers AAV20083-84. The degeneracy of the PCR
XX primers produced several different sequences, of which the present
XX sequence is a consensus sequence. The present sequence was used to

```

```

CC construct chimeric, humanised Act-1 antibodies, which contain murine
CC antigen binding regions. The humanised immunoglobulin can be used to
CC inhibit the interaction of cells bearing alpha4-beta7 with cells bearing
CC a ligand for alpha4-beta7. It can be used for inhibiting leukocyte
CC infiltration of tissues, e.g. for treating inflammatory diseases such
CC as inflammatory bowel disease. The immunoglobulin can also be used for
CC detection, isolation and diagnosis.
XX
SQ Sequence 535 BP; 126 A; 128 C; 132 G; 149 T; 0 other;

Alignment Scores:
Pred. No.:      2.14e-19      Length:      535
Score:          31.00        Matches:      31
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    21.38%      Indels:      0
DB:             19          Gaps:      0

US-08-836-455-2 (1-145) x AAV20086 (1-535)

QY 115 ProTyrThrPheGlyGlyThrLysLeuGluIleLysArgAlaAspAlaAProThr 134
DB 373 CCGTACACGTTCCGAGGGGGACCAAGCTGGAAATAAACGGGCTGATGTCACCAACT 432

QY 135 ValSerIlePheProProSerSerLysLeuGly 145
DB 433 GTATCCATCTTCCACCATCCAGTAAGCTTGGG 465

RESULT 7
AAQ90431
ID AAQ90431 standard; DNA; 438 BP.
XX
AC AAQ90431;
XX
DT 02-FEB-1996 (first entry)
XX
DE DNA encoding anti-idiotypic antibody Idi017 clone 17Kb1.
XX
KW Antibody; cancer; CDR; heavy chain; light chain; immunoglobulin;
KW complementarity determining region, ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 1..438
FT FT /*tag= a
FT FT /product= anti-idiotypic_antibody_Idi017
FT sig_peptide 1..39
FT FT /*tag= b
XX
XX JP07101999-A.
XX
XX 18-APR-1995.
XX
XX 06-OCT-1993; 93JP-0272950.
XX
XX 06-OCT-1993; 93JP-0272950.
XX
XX (HAGI/) HAGIWARA Y.
XX
XX WPI; 1995-182987/24.
XX P-PSDB; AAR74966.
XX
XX Novel anti-idiotypic antibody against an human anticancer monoclonal
XX antibody - and DNA sequences encoding the antibody, useful in
XX pharmacology, medicine and biochemical fields.
XX
XX Example 5; Page 19; 28pp; Japanese.
XX
XX AAQ90425-Q90434 are DNA molecules encoding anti-idiotypic antibodies
XX Idi03, Idi017, Idi020, Idi027 and Idi033 against a human anticancer
XX monoclonal antibody. These antibodies and DNA encoding them are useful
XX in pharmacological, medical and biochemical fields of research.

```

```

XX SQ Sequence 438 BP; 104 A; 120 C; 112 G; 102 T; 0 other;
Alignment Scores:
Pred. No.: 1.48e-18 Length: 438
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.69% Indels: 0
DB: 16 Gaps: 0

US-08-836-455-2 (1-145) x AAQ90431 (1-438)
Qy 116 TyrThrPheGlyGlyThrLysLeuGluLeuLysArgAlaAspAlaProThrVal 135
Db 334 TACACGTTCCGAGGGGACCAAGCTGGAATAAAACGGGCTGATGCTGCACCAACTGTA 393
Qy 136 SerTlePheProSerSerLysLeuGly 145
Db 394 TCCATCTTCCCAACCATCCAGTAAGCTTGGG 423

RESULT 8
AAX00879
ID AAX00879 standard; DNA; 276 BP.
XX AC AAX00879;
XX DT 29-MAR-1999 (first entry)
XX DE Mouse derived RT3 phase antibody light chain pattern C genetic sequence.
XX KW Catalytic; antibody; phase display; immunising; phase expression vector;
XX KW prodrug; scFv; ss.
XX OS Mus sp.
XX FH Key Location/Qualifiers
XX FT CDS 1..276
XX FT /*tag= a
XX FT /note= "the start and stop codons are not indicated"
XX PN US5855885-A.
XX XX 05-JAN-1999.
XX PD 14-JUL-1994; 94US-0273146.
XX PF 22-JAN-1993; 93US-0007684.
XX PR 14-JUL-1994; 94US-0273146.
XX XX (CHIS/) CHISWELL D.
XX PA (DARS/) DARSLEY M J.
XX PA (FITZ/) FITZGERALD K.
XX PA (KENT/) KENTEN J H.
XX PA (MART/) MARTIN M T.
XX PA (MCCA/) MCCAFFERTY J.
XX PA (SMIT/) SMITH R.
XX PA (TITM/) TITMAS R C.
XX PA (WILL/) WILLIAMS R O.
XX XX Chiswell D, Darsley MJ, Fitzgerald K, Kenten JH;
XX PI Martin MT, Mcafferty J, Smith R, Titmas RC, Williams RO;
XX XX WPI; 1999-105036/09.
XX DR P-FSDB; AAW95480.
XX XX Production of catalytic antibodies displayed on bacteriophages -
XX PT comprises generating a gene library of antibody-derived domains
XX PT inserting coding into a phase expression vector and isolating the
XX PT catalytic antibodies
XX XX Example 4; Fig 11; 117pp; English.

```

---

```

CC The invention relates to methods for producing catalytic antibodies
CC displayed on a phage. The method comprises: (a) generating a gene library
CC of antibody-derived domains; (b) inserting coding for the domains into a
CC phase expression vector; and (c) isolating the catalytic antibodies. The
CC phase expression vector incorporates a histidine peptide in tandem with a
CC myc peptide. The catalytic antibodies can be isolated by preparing an
CC antigen; optionally immunising an animal with the antigen; generating a
CC library of VH and VL domains from the immunised animal; cloning the VH
CC and VL domains into a phase expression vector to generate phage display
CC antibodies; selecting phage display antibodies which bind specifically
CC to the antigen; screening the selected phage display antibodies for
CC catalytic activity to substrate; and isolating the catalytic antibodies,
CC where the phase expression vector incorporates a histidine peptide in
CC tandem with a myc peptide. The processes are used to produce catalytic
CC antibodies, which can be used for in vivo activation of a prodrug. The
CC present sequence represents a genetic sequence of light chain PCR
CC pattern C from mouse derived RT3 phase antibodies.
XX SQ Sequence 276 BP; 67 A; 65 C; 64 G; 80 T; 0 other;

Alignment Scores:
Pred. No.: 8.01e-18 Length: 276
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 20 Gaps: 0

US-08-836-455-2 (1-145) x AAX00879 (1-276)
Qy 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95
Db 142 TCTGTGTCCTCCCAAGGTTTCAGTGGCAGTAGTCTGGGTGAGATTCTCTCACCATC 201
Qy 96 SerSerLeuGluSerGluAspPheVal 104
Db 202 AGCAGCCTTCAGTCTGAGATTTCGTA 228

RESULT 9
AAX00875
ID AAX00875 standard; DNA; 276 BP.
XX AC AAX00875;
XX DT 29-MAR-1999 (first entry)
XX DE Mouse derived RT3 phase antibody light chain pattern A genetic sequence.
XX KW Catalytic; antibody; phase display; immunising; phase expression vector;
XX KW prodrug; scFv; ss.
XX OS Mus sp.
XX FH Key Location/Qualifiers
XX FT CDS 1..276
XX FT /*tag= a
XX FT /note= "the start and stop codons are not indicated"
XX PN US5855885-A.
XX XX 05-JAN-1999.
XX PD 14-JUL-1994; 94US-0273146.
XX PF 22-JAN-1993; 93US-0007684.
XX PR 14-JUL-1994; 94US-0273146.
XX XX (CHIS/) CHISWELL D.
XX PA (DARS/) DARSLEY M J.
XX PA (FITZ/) FITZGERALD K.
XX PA (KENT/) KENTEN J H.
XX PA (MART/) MARTIN M T.
XX PA (MCCA/) MCCAFFERTY J.
XX PA (SMIT/) SMITH R.
XX PA (TITM/) TITMAS R C.
XX PA (WILL/) WILLIAMS R O.
XX XX Chiswell D, Darsley MJ, Fitzgerald K, Kenten JH;
XX PI Martin MT, Mcafferty J, Smith R, Titmas RC, Williams RO;
XX XX WPI; 1999-105036/09.
XX DR P-FSDB; AAW95480.
XX XX Production of catalytic antibodies displayed on bacteriophages -
XX PT comprises generating a gene library of antibody-derived domains
XX PT inserting coding into a phase expression vector and isolating the
XX PT catalytic antibodies
XX XX Example 4; Fig 11; 117pp; English.

```

PA (SMIT/) SMITH R.  
PA (TITM/) TITMAS R C.  
PA (WILL/) WILLIAMS R O.  
XX Chiswell D, Darsley MJ, Fitzgerald K, Kenten JH;  
PI Martin MT, McCafferty J, Smith R, Titmas RC, Williams RO;  
XX WPI; 1999-105036/09.  
DR P-PSDB; AAW95476.  
XX Production of catalytic antibodies displayed on bacteriophages -  
PT comprises generating a gene library of antibody-derived domains  
PT inserting coding into a phage expression vector and isolating the  
PT catalytic antibodies  
XX Example 4; Fig 9A-F; 117pp; English.  
XX The invention relates to methods for producing catalytic antibodies  
CC displayed on a phage. The method comprises: (a) generating a gene library  
CC of antibody-derived domains; (b) inserting coding for the domains into a  
CC phage expression vector; and (c) isolating the catalytic antibodies. The  
CC phage expression vector incorporates a histidine peptide in tandem with a  
CC myc peptide. The catalytic antibodies can be isolated by preparing an  
CC antigen; optionally immunising an animal with the antigen; generating a  
CC library of VH and VL domains from the immunised animal; cloning the VH  
CC and VL domains into a phage expression vector to generate phage display  
CC antibodies; selecting phage display antibodies which bind specifically  
CC to the antigen; screening the selected phage display antibodies for  
CC catalytic activity to substrate; and isolating the catalytic antibodies  
CC where the phage expression vector incorporates a histidine peptide in  
CC tandem with a myc peptide. The processes are used to produce catalytic  
CC antibodies, which can be used for in vivo activation of a prodrug. The  
CC present sequence represents a genetic sequence of light chain pattern A  
CC from mouse derived RT3 phage antibodies.  
XX SQ Sequence 276 BP; 68 A; 65 C; 53 G; 80 T; 0 other;

Alignment Scores:  
Pred. No.: 8,01e-18 Length: 276  
Score: 29.00 Matches: 29  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.00% Indels: 0  
DB: 20 Gaps: 0

US-08-836-455-2 (1-145) x AAX00875 (1-276)  
QY 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95  
Db 142 TCTGGTGTCCCAAGAGGTTTCAGTGGCAGTAGGTCTGGTGCAGATTATTCTCACCATC 201  
QY 96 SerSerLeuGluSerGluAspPheVal 104  
Db 202 AGCAGCCTTCAGTCTGAAGATTGTGA 228

RESULT 10  
AAA38908  
ID AAA38908 standard; DNA; 321 BP.  
XX AC AAA38908;  
XX AC  
XX DT 29-AUG-2000 (first entry)  
XX DE 520C9 hybridoma VL domain encoding DNA SEQ ID NO:25.  
XX Antigen binding site; immunoglobulin; cancer antigen; immunological;  
KW antibody; tumour; human; mucin; cancer; cytostatic; hybridoma;  
KW specific binding assay; affinity purification; drug targeting;  
KW toxin targeting; imaging; genetic; therapeutic; ss.  
XX Homo sapiens.  
XX OS  
XX PH US6054561-A.

XX 25-APR-2000.  
XX 07-JUN-1995; 95US-0483749.  
XX 21-MAR-1986; 86US-0842476.  
PR 08-MAY-1988; 88US-0190778.  
PR 08-FEB-1984; 84US-0577976.  
PR 11-JAN-1985; 85US-0690750.  
PR 11-AUG-1994; 94US-0288981.  
XX (CHIR ) CHIRON CORP.  
XX PI Ring DB;  
XX WPI; 2000-338508/29.  
DR P-PSDB; AAY90824.  
XX Monoclonal antibody capable of binding to human breast cancer antigen  
PT useful for affinity purification, drug or toxin targeting, imaging, and  
PT treating cancer  
XX Disclosure; Fig 13; 57pp; English.  
XX The present invention describes a monoclonal antibody (MAB) (I) that  
CC binds to a human breast cancer antigen that is also bound by MAB 454C11  
CC and 520C9 (produced hybridoma ATCC HB8484 and HB8696, respectively). Also  
CC described is a hybridoma that produces (I). (I) is useful in specific  
CC binding assays, affinity purification, drug or toxin targeting, imaging,  
CC and genetic or immunological therapeutics for various cancers. The  
CC present sequence encodes a VL domain derived from a 520C9 hybridoma,  
CC which is used in the exemplification of the present invention.  
XX SQ Sequence 321 BP; 82 A; 78 C; 72 G; 89 T; 0 other;

Alignment Scores:  
Pred. No.: 9,24e-18 Length: 321  
Score: 29.00 Matches: 29  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.00% Indels: 0  
DB: 21 Gaps: 0

US-08-836-455-2 (1-145) x AAA38908 (1-321)  
QY 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95  
Db 166 TCTGGTGTCCCAAGAGGTTTCAGTGGCAGTAGGTCTGGTGCAGATTATTCTCACCATC 225  
QY 96 SerSerLeuGluSerGluAspPheVal 104  
Db 226 AGTAGCCTTCAGTCTGAAGATTGTGA 252

RESULT 11  
AAZ49534  
ID AAZ49534 standard; cDNA; 324 BP.  
XX AC AAZ49534;  
XX AC  
XX DT 04-APR-2000 (first entry)  
XX DE Mouse anti-IL-18 antibody VL encoding cDNA from hybridoma #125-2H.  
XX Anti-Interleukin-18 antibody; IL-18; light chain variable region; VL;  
KW hybridoma #125-2H; mouse; monoclonal antibody #125-2HmAb; interleukin-18;  
KW antiinflammatory; immunosuppressive; leucocytopenic; antialgic;  
KW antipyretic; antiallergic; hepatotropic; treatment; autoimmune disease;  
KW immunopathy; inflammatory disorder; immunoreaction; ss.  
XX Mus musculus.  
XX OS  
XX PH Key  
XX mat\_peptide Location/Qualifiers  
XX 1..324

```
FT      /*tag= a
FT      /label= Anti-IL-18_antibody_light_chain_variable_region
PN      EP974600-A2.
XX
XX      PD      26-JAN-2000.
XX
XX      PF      24-JUN-1999; 99EP-0304977.
XX
XX      PR      24-JUN-1998; 98JP-0177580.
XX      PR      12-OCT-1998; 98JP-0289044.
XX      PR      22-DEC-1998; 98JP-0385023.
XX
XX      PA      (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX      PI      Nishida Y, Okura T, Tanimoto T, Kurimoto M;
XX
XX      DR      WPI; 2000-118341/11.
XX      DR      P-PSDB; AAY44587.
XX
XX      PT      New artificially produced peptide for neutralizing biological activity
XX      PT      of interleukin-18, useful for treating and preventing immunopathies,
XX      PT      inflammatory disorders and autoimmune diseases.
XX
XX      PS      Claim 11; Page 21; 36pp; English.
XX
XX      CC      The present cDNA sequence derived from hybridoma #125-2H, encodes mouse
XX      CC      anti-interleukin-18 antibody light chain variable region (VL). It can be
XX      CC      used in the production of recombinant monoclonal antibodies #125-2HmAb,
XX      CC      which is capable of neutralising biological activities of interleukin-18.
XX      CC      The antibody has antiinflammatory, immunosuppressive, leucocytopenetic,
XX      CC      antiangiogenic, antipretic, antiallergic and hepatotropic activity and can
XX      CC      be used for prevention and treatment of autoimmune diseases, immunopathies
XX      CC      and inflammatory disorders caused by excessive immunoreaction.
XX
XX      SQ      Sequence 324 BP; 83 A; 78 C; 76 G; 87 T; 0 other;

Alignment Scores:
Pred. No.:      9.32e-18      Length:      324
Score:      29.00      Matches:      29
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      20.00%      Indels:      0
DB:      21      Gaps:      0

US-08-836-455-2 (1-145) x AAZ49534 (1-324)

Oy      76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95
Db      166 TCTGGTGTCCCAAGAGGTTTCAGTGGCAGTAGGTCAGTGGTCAGATTATCTCTCACCATC 225

Oy      96 SerSerLeuGluSerGluAspPheVal 104
Db      226 AGCAGCCTTGAGTCTGAAGATTGTGA 252

RESULT 12
AAZ52607
ID      AAZ52607 standard; cDNA; 390 BP.
XX
XX      AC      AAZ52607;
XX
XX      DT      14-MAY-2003 (first entry)
XX
XX      DE      Escherichia coli light chain variable region cDNA.
XX
XX      KW      S-surface antigen; Hepatitis B virus; HBV; therapy; infection; virucide;
XX      KW      hepatotropic; gene; ss.
XX      OS      Escherichia coli.
XX
XX      Key      Location/Qualifiers
FT      CDS      1..390
FT      /*tag= a
```

```
FT      /product= "Escherichia coli light chain variable region"
FT      /note= "CDS does not include start and stop codon"
FT      /partial
PN      WO200292819-Al.
XX
XX      PD      21-NOV-2002.
XX
XX      PF      15-MAY-2002; 2002WO-KR00905.
XX
XX      PR      16-MAY-2001; 2001KR-0026634.
XX
XX      PA      (YUHA-) YUHAN CORP.
XX
XX      PI      Lee JW, Ko IY, Kang HK, Song MY, Song TH, Kim CS;
XX
XX      DR      WPI; 2003-140281/13.
XX      DR      P-PSDB; AAE34366.
XX
XX      PT      New light and heavy chain variable regions of a monoclonal antibody
XX      PT      against the S-surface antigen of the hepatitis B virus (HBV), useful
XX      PT      for neutralizing or removing HBV, or for preventing or treating HBV
XX      PT      infection.
XX
XX      PS      Claim 3; Page 15; 20pp; English.
XX
XX      CC      The invention relates to light and heavy chain variable regions of a
XX      CC      monoclonal antibody against S-surface antigen of the hepatitis B virus
XX      CC      (HBV). The variable regions of the antibodies are useful against HBV
XX      CC      S-surface antigens, e.g. adr, adw, ayr or ayw, particularly for
XX      CC      neutralising or removing HBV. They may also be employed to treat or
XX      CC      prevent HBV infection. The present sequence is Escherichia coli light
XX      CC      chain variable region cDNA.
XX
XX      SQ      Sequence 390 BP; 99 A; 96 C; 98 G; 97 T; 0 other;

Alignment Scores:
Pred. No.:      1.11e-17      Length:      390
Score:      29.00      Matches:      29
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      20.00%      Indels:      0
DB:      25      Gaps:      0

US-08-836-455-2 (1-145) x AAD52607 (1-390)

Oy      117 ThrPheGlyGlyGlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSer 136
Db      304 ACGTTCGGTGGAGGCCACCACAGCTGGAATCAACAGGGCTGATGCTGCACCACTGTATCC 363

Oy      137 IlePheProProSerSerLysLeuGly 145
Db      364 ATCTTCCACCACCATCCAGTAAGCTTGGG 390

RESULT 13
AAZ49548
ID      AAZ49548 standard; cDNA; 407 BP.
XX
XX      AC      AAZ49548;
XX
XX      DT      04-APR-2000 (first entry)
XX
XX      DE      Mouse light chain variable region DNA.
XX
XX      KW      Mouse light chain variable region; VL; variable region light chain;
XX      KW      interleukin-18; immunopathies; inflammatory disorder; autoimmune disease;
XX      KW      mouse; anti-allergic; anti-inflammatory; immunosuppressive;
XX      KW      hematopoietic; leukocytopenetic; antiangiogenic; antipretic; ds.
XX      OS      Mus musculus.
XX      OS      Synthetic.
XX
XX      Key      Location/Qualifiers
FT
```



```

FT mat_peptide 1..407
FT /*tag= a
FT /product= "Mouse light chain variable region"
FT sig_peptide 1..60
FT /*tag= b
XX
PN EP974600-A2.
XX
PD 26-JAN-2000.
XX
PE 24-JUN-1999; 99EP-0304977.
XX
PR 24-JUN-1998; 98JP-0177580.
PR 12-OCT-1998; 98JP-0289044.
PR 22-DEC-1998; 98JP-0365023.
XX
PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Nishida Y, Okura T, Tanimoto T, Kurimoto M;
XX WPI; 2000-118341/11.
DR P-PSDB; AAY44599.
XX
PT New artificially produced peptide for neutralizing biological activity
PT of interleukin-18, useful for treating and preventing immunopathies,
PT inflammatory disorders and autoimmune diseases -
XX
PS Example 1; Page 28-29; 32pp; English.
XX
CC The present sequence encodes mouse light chain variable region. This
CC recombinant DNA is derived from PCR A which amplifies antibody light
CC chain variable region (VL). The transformant produced using the VL gene
CC was used transform competent E. coli cells. The peptide produced by
CC transformants neutralizes interleukin-18. This is useful for treating and
CC preventing immunopathies, inflammatory disorders and autoimmune diseases
CC which are caused by excessive immunoreaction. The peptide has
CC anti-allergic, anti-inflammatory, immunosuppressive, hematopoietic,
CC leukocytopenic, antialgic, antipyretic and hepatic-function improving
CC activities.
XX
SQ Sequence 407 BP; 96 A; 99 C; 96 G; 116 T; 0 other;

Alignment Scores:
Pred. No.: 1.16e-17 Length: 407
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 21 Gaps: 0

US-08-836-455-2 (1-145) x AA249548 (1-407)

QY 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95
Db 226 TCTGGTGTCCCAAGAGGTTCACTGGCAGTAGCTGGGTCAGATTATCTCCACCATC 285
QY 96 SerSerLeuGluSerGluAspPheVal 104
Db 286 AGCAGCGCTTGAGTCTCAAGATTGTGA 312

RESULT 14
AAT31332
ID AAT31332 standard; cDNA; 447 BP.
XX
AC AAT31332;
XX
XX 25-MAR-2003 (updated)
DT 26-FEB-1997 (first entry)
XX
XX Anti-Idiotypic monoclonal antibody 1A7 variable light chain, cDNA.
DE
XX Murine; mouse; anti-idiotype; monoclonal antibody; MAb; 1A7;
XX variable light chain; ganglioside 2; GD2; 14G2a; neuroblastoma;
KW

```

```

KW glycosphingolipid; human; neuroectodermal; tumour; glioma; lung;
KW malignant melanoma; soft tissue sarcoma; small cell carcinoma;
KW vaccine; treatment; palliate; detection; diagnosis;
KW recombinant production; purification; probe; primer; assay;
KW amplification; gene therapy; ss.
XX
OS Mus musculus.
XX
XX
FH Key Location/Qualifiers
FT mat_peptide 1..446
FT /*tag= a
XX
XX
PN WO9622373-A2.
XX
PD 25-JUL-1996.
XX
XX
PF 17-JAN-1996; 96WO-US00882.
XX
XX
PR 17-JAN-1995; 95US-0372676.
PR 16-JAN-1996; 96US-0591196.
PR 17-JAN-1995; 95US-0372676.
PR 16-JAN-1996; 96US-0591196.
XX
PA (KENT ) UNIV KENTUCKY.
XX
XX Chatterjee M, Chatterjee SK, Foon KA;
XX
DR WPI; 1996-354530/35.
DR P-PSDB; AAW03199.
XX
XX Monoclonal antibody 1A7 and related polynucleotide(s) and
XX polyptide(s) - useful to treat or palliate a GD2-associated
XX disease, e.g. melanoma and glioma
XX
PS Claim 10; Fig 1; 14pp; English.
XX
CC The present sequence encodes the murine anti-idiotype monoclonal
CC antibody (MAb) 1A7 variable light chain. MAb 1A7 was raised against
CC the anti-ganglioside 2 (GD2) MAb 14G2a, which binds an unique
CC epitope of GD2. As the glycosphingolipid GD2 is expressed at high
CC density by human neuroectodermal tumours, e.g. malignant melanoma,
CC neuroblastoma, glioma, soft tissue sarcoma and small cell carcinoma
CC of the lung, MAb 1A7, or its cDNA can be used in a vaccine to treat
CC or palliate such diseases. They can also be used to reduce the
CC risk of recurrence of a clinically detectable tumour, and detect an
CC anti-GD2 Ab bound to a tumour cell.
CC MAb 1A7 overcomes immune tolerance and induces an immune response
CC against GD2, which comprises anti-GD2 Ab (humoral response) and
CC GD2-specific cells (cellular response). It can be used to purify
CC anti-1A7 (Ab3), anti-GD2 (Ab1') or 14G2a (Ab1), detect anti-1A7 or
CC anti-GD2 in a sample or measure the level of cellular anti-1A7 or
CC anti-GD2 activity.
CC The cDNA can be used in expression systems for 1A7 prodn., and in
CC the prepn. of probes and primers to respectively assay for 1A7
CC cDNA, and amplify desired polynucleotides for use in gene therapy.
CC (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 447 BP; 106 A; 110 C; 108 G; 123 T; 0 other;

Alignment Scores:
Pred. No.: 1.27e-17 Length: 447
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 17 Gaps: 0

US-08-836-455-2 (1-145) x AAT31332 (1-447)

QY 117 ThrPheGlyGlyThrLysLeuGluLeuLysArgAlaAspAlaProThrValSer 136
Db 361 ACGTTCGGTGGAGGCACCAAGCTGGGAANTCAACAGGCGCTGCTGCCACCACTGTATCC 420

```

Search completed: August 31, 2003, 01:16:24  
Job time : 218.987 secs

```

Qy 137 IlePheProProSerSerLysLeuGly 145
Db 421 ATCTTCCACCATCCAGTAGCTTGGG 447

RESULT 15
AAZ31365
ID AAZ31365 standard; cDNA; 447 BP.
XX
AC AAZ31365;
XX
DT 07-FEB-2000 (first entry)
XX
DE MAB 1A7 light chain variable region encoding cDNA.
XX
KW Monoclonal antibody; MAB; 1A7; GD2; immune response; melanoma;
KW neuroblastoma; glioma; soft tissue carcinoma; small cell carcinoma;
KW tumor-associated antigen; ss.
XX
OS Synthetic.
OS Mus sp.
XX
FN US9777316-A.
XX
PD 02-NOV-1999.
XX
PF 16-JAN-1996; 96US-0591196.
XX
PR 17-JAN-1995; 95US-0372676.
XX
PA (KENT ) UNIV KENTUCKY.
XX
PI Foon KA, Chatterjee SK, Chatterjee M;
XX
DR WPI; 1999-619711/53.
XX
DR P-PSDB; AAY49209.
XX
PT Monoclonal antibody 1A7 which elicits an anti-GD2 immunological
PT response, useful for the development of products for the detection and
PT treatment of cancers .
XX
PS Disclosure; Fig 1; 74pp; English.
XX
CC The invention provides a monoclonal antibody (MAB) designated 1A7, which
CC elicits an anti-GD2 (tumor-associated antigen) immunological response in
CC humans. MAB 1A7 has defined light and heavy chain variable region
CC sequences. The MAB 1A7 and polypeptides can be used for eliciting an
CC anti-GD2 immune response. The polypeptides can also be used for detecting
CC or purifying anti-GD2 antibody. The products can be used for treating GD2
CC -associated diseases, e.g. melanoma, neuroblastoma, glioma, soft tissue
CC carcinoma, and small cell carcinoma. They can be used for palliating the
CC disease or for reducing the risk of recurrence. The present sequence
CC represents the cDNA encoding the light chain variable region of MAB 1A7.
XX
SQ Sequence 447 BP; 106 A; 111 C; 107 G; 123 T; 0 Other;

Alignment Scores:
Pred. NO.: 1.27e-17 Length: 447
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 20 Gaps: 0

US-08-836-455-2 (1-145) x AAZ31365 (1-447)

Qy 117 ThrPheGlyGlyThrLysLeuGlyAlaAspAlaProThrValSer 136
Db 361 ACGTTCGGTGAGCCACCAAGCTGGAATCAACGGGCTGATGCTGCCAACCTGTATCC 420

Qy 137 IlePheProProSerSerLysLeuGly 145
Db 421 ATCTTCCACCATCCAGTAGCTTGGG 447

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 31, 2003, 01:02:06 ; Search time 192.198 Seconds  
(without alignments)  
1730.853 Million cell updates/sec

Title: US-08-836-455-2  
Perfect score: 145  
Sequence: 1 MGAQAIGLFLLLPGTRC.....IKRADAAPTSTIFFPSSKIG 145

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1533700 segs, 1147125425 residues  
Word size: 1

Total number of hits satisfying chosen parameters: 3063161  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q/cgn2.1/USPTO.spool/US08836455/runat\_29082003\_132954\_23062/app\_query.fasta\_1.654  
-DB=PublishedApplications.NA -QSWT=fastap -SUFFIX=p2noli.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNIQS=bits -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US08836455.ecgn\_1\_1\_271\_runat\_29082003\_132954\_23062  
-NCPU=6 -ICPU=3 -NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PublishedApplications.NA:

- 1: /cgn2.6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2.6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2.6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2.6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2.6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2.6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2.6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2.6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2.6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2.6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2.6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2.6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2.6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2.6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2.6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2.6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2.6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	145	100.0	435	9 US-09-861-294-1 Sequence 1, Appli

2	145	100.0	435	12 US-10-367-506-1	Sequence 1, Appli
3	33	22.8	387	9 US-09-007-093-1	Sequence 1, Appli
4	29	20.0	324	10 US-09-924-099-11	Sequence 11, Appl
5	29	20.0	351	8 US-08-779-784-16	Sequence 16, Appl
6	29	20.0	351	8 US-08-779-784-17	Sequence 17, Appl
7	29	20.0	390	13 US-10-146-305-5	Sequence 5, Appli
8	29	20.0	402	8 US-08-779-784-5	Sequence 27, Appl
9	29	20.0	407	10 US-09-924-099-27	Sequence 1, Appli
10	29	20.0	447	10 US-09-990-205-1	Sequence 1, Appli
11	29	20.0	447	14 US-10-153-401-1	Sequence 19, Appl
12	29	20.0	711	10 US-09-924-099-19	Sequence 20, Appl
13	29	20.0	729	10 US-09-924-099-20	Sequence 5, Appli
14	29	20.0	739	10 US-09-887-853-5	Sequence 10, Appl
15	28	19.3	390	12 US-10-268-883-10	Sequence 18, Appl
16	28	19.3	698	13 US-10-006-773-18	Sequence 95, Appl
17	26	17.9	426	10 US-09-840-459-95	Sequence 101, App
18	26	17.9	426	10 US-09-840-459-101	Sequence 22, Appl
19	26	17.9	705	12 US-10-281-479A-22	Sequence 10, Appl
20	26	17.9	714	14 US-10-216-484-10	Sequence 10, Appl
21	26	17.9	729	13 US-10-006-773-10	Sequence 3, Appli
22	26	17.9	831	10 US-09-903-327A-3	Sequence 8, Appli
23	26	17.9	5711	10 US-09-897-006-8	Sequence 13, Appl
24	26	17.9	5711	11 US-09-897-511A-8	Sequence 13, Appl
25	26	17.9	6255	11 US-09-897-006-13	Sequence 1, Appli
26	26	17.2	6255	11 US-09-897-511A-13	Sequence 3, Appli
27	25	17.2	447	9 US-09-797-481-1	Sequence 19, Appl
28	25	17.2	447	9 US-09-861-294-19	Sequence 3, Appli
29	25	17.2	447	12 US-10-367-506-19	Sequence 3, Appli
30	25	17.2	447	14 US-10-162-396-3	Sequence 14, Appl
31	25	17.2	447	14 US-10-006-773-14	Sequence 3, Appli
32	25	17.2	504	13 US-09-469-485-3	Sequence 9, Appli
33	21	14.5	351	11 US-09-929-665-9	Sequence 10, Appl
34	21	14.5	363	11 US-09-929-665-10	Sequence 9, Appli
35	21	14.5	363	11 US-09-929-546-9	Sequence 10, Appl
36	21	14.5	363	11 US-09-929-546-10	Sequence 10, Appl
37	21	14.5	363	11 US-09-967-719C-3	Sequence 1, Appli
38	21	14.5	809	13 US-10-027-770-1	Sequence 20, Appl
39	20	13.8	384	9 US-09-905-243-20	Sequence 7, Appli
40	19	13.1	737	10 US-09-905-243-27	Sequence 103, App
41	18	12.4	387	9 US-09-940-727B-103	Sequence 37859, A
42	17	11.7	368	11 US-09-918-995-37859	Sequence 2025, Ap
43	17	11.7	490	11 US-10-066-543-2025	
44	17	11.7	514		
45	17	11.7			

ALIGNMENTS

RESULT 1  
US-09-861-294-1  
; Sequence 1, Application US/09861294  
; Patent No. US20020098190A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; FILE OF INVENTION: TUMORS BEARING HMF AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 09/096,244  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 435  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(435)

```
; NAME/KEY: sig_peptide
; LOCATION: (1)...(60)
; NAME/KEY: mat_peptide
; LOCATION: (61)...(435)
US-09-861-294-1

Alignment Scores:
Pred. No.: 2,36e-131 Length: 435
Score: 145.00 Matches: 145
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-08-836-455-2 (1-145) x US-09-861-294-1 (1-435)

Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20
Db 1 ATGGGGGCCCTGCTCAGATTCTTGGGTTCTTGTGCTCTTGTTCAGGTACCAGATGT 60
Qy 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40
Db 61 GACATCCAGATGACCCAGTCTCCATCTCTATCTGCTCTCTGGGACAAAGAGTCACT 120
Qy 41 LeuThrCysArgAlaSerGlnAspileGlyIleAsnLeuHisTrpLeuGlnGluPro 60
Db 121 CTCACCTTGTTCGGGCAAGTCAGGACATGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180
Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerLeuGlySerGlyValProLys 80
Db 181 GATGGAACTATTAAACCCCTGATCTACGCCACATCCAGTTAGGTTCTGGTGTCCCCAAA 240
Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100
Db 241 AGGTTACAGTGGCAGTAGTCTGGGTCAGATTATTTCTCTCACCATCAGCAGCTTGAGTCT 300
Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerProTyrThrPheGlyGly 120
Db 301 GAAGATTTTGTAGCCTATTACTGTCTACAATATGCTTCTCGGTACACGTTTCGGAGGG 360
Qy 121 GlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSerIlePheProPro 140
Db 361 GGGACCAAGCTGGAAATAAACGGCGTGTGCTGCACCAACTGTATCCATCTTCCCCACCA 420
Qy 141 SerSerLysLeuGly 145
Db 421 TCCAGTAAGCTTGGG 435

RESULT 2
US-10-367-506-1
; Sequence 1, Application US/10367506
; Publication No. US20030152575A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth A. FOON
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; FILE OF INVENTION: TUMORS BEARING HMGF AND CEA ANTIGENS
; FILE REFERENCE: 304142000620
; CURRENT APPLICATION NUMBER: US/10/367,506
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US/09/861,294
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/049,540
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 09/096,244
; PRIOR FILING DATE: 1998-06-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:

; NAME/KEY: CDS
; LOCATION: (1)...(435)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (1)...(60)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (61)...(435)
US-10-367-506-1

Alignment Scores:
Pred. No.: 2,36e-131 Length: 435
Score: 145.00 Matches: 145
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-08-836-455-2 (1-145) x US-10-367-506-1 (1-435)

Qy 1 MetGlyAlaProAlaGlnIleLeuGlyPheLeuLeuLeuPheProGlyThrArgCys 20
Db 1 ATGGGGGCCCTGCTCAGATTCTTGGGTTCTTGTGCTCTTGTTCAGGTACCAGATGT 60
Qy 21 AspileGlnMetThrGlnSerProSerSerLeuSerAlaSerLeuGlyGlnArgValSer 40
Db 61 GACATCCAGATGACCCAGTCTCCATCTCTATCTGCTCTCTGGGACAAAGAGTCACT 120
Qy 41 LeuThrCysArgAlaSerGlnAspileGlyIleAsnLeuHisTrpLeuGlnGluPro 60
Db 121 CTCACCTTGTTCGGGCAAGTCAGGACATGGTATTAACTTACATTTGGCTTCAGCAGGAACCA 180
Qy 61 AspGlyThrIleLysArgLeuIleTyrAlaThrSerLeuGlySerGlyValProLys 80
Db 181 GATGGAACTATTAAACCCCTGATCTACGCCACATCCAGTTAGGTTCTGGTGTCCCCAAA 240
Qy 81 ArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIleSerSerLeuGluSer 100
Db 241 AGGTTACAGTGGCAGTAGTCTGGGTCAGATTATTTCTCTCACCATCAGCAGCTTGAGTCT 300
Qy 101 GluAspPheValAlaTyrTyrCysLeuGlnTyrAlaSerProTyrThrPheGlyGly 120
Db 301 GAAGATTTTGTAGCCTATTACTGTCTACAATATGCTTCTCGGTACACGTTTCGGAGGG 360
Qy 121 GlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSerIlePheProPro 140
Db 361 GGGACCAAGCTGGAAATAAACGGCGTGTGCTGCACCAACTGTATCCATCTTCCCCACCA 420
Qy 141 SerSerLysLeuGly 145
Db 421 TCCAGTAAGCTTGGG 435

RESULT 3
US-09-007-093-1
; Sequence 1, Application US/09007093
; Patent No. US20020025315A1
; GENERAL INFORMATION:
; APPLICANT: Anand, Naveen N
; APPLICANT: Barber, Brian H
; APPLICANT: Cates, George A
; APPLICANT: Caterini, Judith E
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF
; NUMBER OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/007,093
; APPLICATION NUMBER:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,576
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-765
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-09-007-093-1
;
Alignment Scores:
Pred. No.: 5,28e-23 Length: 387
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.76% Indels: 0
DB: 9 Gaps: 0

US-08-836-455-2 (1-145) x US-09-007-093-1 (1-387)

QY 15 PheProGlyThrArgCysAspIleGlnMetThrGlnSerProSerSerLeuSerAlaSer 34
Db 49 TTTCCAGGTACCAAGATGTGACATGCCAGTCCAGTCCATCCCTCTATCTGCCTCT 108
QY 35 LeuGlyGlnArgValSerLeuThrCysArgAlaSerGln 47
Db 109 CTGGGACAAGAGTCACTCTCACTTGTGGGCAAGTCAG 147

RESULT 4
US-09-924-099-11
; Sequence 11, Application US/09924099
; Patent No. US20020128450A1
; GENERAL INFORMATION:
; APPLICANT: NISHIDA, Yoshihiro
; APPLICANT: OKURA, Takanori
; APPLICANT: TANIMOTO, Tadao
; APPLICANT: KURIMOTO, Masashi
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/924,099
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 11
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Mus musculus
;
US-09-924-099-11
Alignment Scores:
Pred. No.: 3,35e-19 Length: 324
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 8 Gaps: 0

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/007,093
; APPLICATION NUMBER:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,576
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-765
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-09-007-093-1
;
Alignment Scores:
Pred. No.: 5,28e-23 Length: 387
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.76% Indels: 0
DB: 9 Gaps: 0

US-08-836-455-2 (1-145) x US-09-007-093-1 (1-387)

QY 15 PheProGlyThrArgCysAspIleGlnMetThrGlnSerProSerSerLeuSerAlaSer 34
Db 49 TTTCCAGGTACCAAGATGTGACATGCCAGTCCAGTCCATCCCTCTATCTGCCTCT 108
QY 35 LeuGlyGlnArgValSerLeuThrCysArgAlaSerGln 47
Db 109 CTGGGACAAGAGTCACTCTCACTTGTGGGCAAGTCAG 147

RESULT 4
US-09-924-099-11
; Sequence 11, Application US/09924099
; Patent No. US20020128450A1
; GENERAL INFORMATION:
; APPLICANT: NISHIDA, Yoshihiro
; APPLICANT: OKURA, Takanori
; APPLICANT: TANIMOTO, Tadao
; APPLICANT: KURIMOTO, Masashi
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/924,099
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 11
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Mus musculus
;
US-09-924-099-11
Alignment Scores:
Pred. No.: 3,35e-19 Length: 324
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 8 Gaps: 0

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x US-09-924-099-11 (1-324)

QY 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95
Db 166 TCTGGTGTCCCAAGAGGTTCACTGGCAGTAGCTCTGGGTGACATATTCTCTCACCATC 225
QY 96 SerSerLeuGluSerGluAspPheVal 104
Db 226 AGCAGCCTTGAGTCTGAAGATTGTGA 252

RESULT 5
US-08-779-784-16
; Sequence 16, Application US/08779784
; Publication No. US20020164325A1
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Moses
; APPLICANT: Miller, David J.
; APPLICANT: Asakura, Kunihiko
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,784
; FILING DATE: 07-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/692,084
; FILING DATE: 08-AUG-1996
; PRIOR APPLICATION DATA: US 08/236,520
; APPLICATION NUMBER: US 08/236,520
; FILING DATE: 29-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
;
US-08-779-784-16
Alignment Scores:
Pred. No.: 3,6e-19 Length: 351
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 8 Gaps: 0
```

```
US-08-836-455-2 (1-145) x US-08-779-784-16 (1-351)
QY 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerLeuThrIle 95
Db 232 TCTGGTGTGCCAAAGGTTTCAGTGGCAGTAGGTCTGGTCCAGATTATTCTCTCACCATC 291
QY 96 SerSerLeuGluSerGluAspPheVal 104
Db 292 AGCAGCCTTGAGTCTGAAGATTGTGA 318
RESULT 6
US-08-779-784-17
; Sequence 17, Application US/08779784
; Publication No. US20020164325A1
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Moses
; APPLICANT: Miller, David J.
; APPLICANT: Asakura, Kunihiko
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM
; TITLE OF INVENTION: REMYLINATION USING MONOCLONAL AUTOANTIBODIES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,784
; FILING DATE: 07-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/692,084
; FILING DATE: 08-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/236,520
; FILING DATE: 29-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHEetical: NO
US-08-779-784-17
Alignment Scores:
Pred. No.: 3.6e-19 Length: 351
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 8 Gaps: 0
US-08-836-455-2 (1-145) x US-08-779-784-17 (1-351)
QY 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerLeuThrIle 95
```

```
Db 232 TCTGGTGTGCCAAAGGTTTCAGTGGCAGTAGGTCTGGTCCAGATTATTCTCTCACCATC 291
QY 96 SerSerLeuGluSerGluAspPheVal 104
Db 292 AGCAGCCTTGAGTCTGAAGATTGTGA 318
RESULT 7
US-10-146-305-5
; Sequence 5, Application US/10146305
; Publication No. US20020173035A1
; GENERAL INFORMATION:
; APPLICANT: YUHAN CORPORATION
; TITLE OF INVENTION: A VARIABLE REGION OF THE MONOCLONAL ANTIBODY AGAINST THE HBV
; TITLE OF INVENTION: S-SURFACE ANTIGEN AND A GENE ENCODING THE SAME
; FILE REFERENCE: OVI7440
; CURRENT APPLICATION NUMBER: US/10/146,305
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: KR 10-2001-26634
; PRIOR FILING DATE: 2001-05-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Kopatentin 1.71
; SEQ ID NO 5
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-146-305-5
Alignment Scores:
Pred. No.: 3.96e-19 Length: 390
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 13 Gaps: 0
US-08-836-455-2 (1-145) x US-10-146-305-5 (1-390)
QY 117 ThrPheGlyGlyThrLysLeuGluIleLysArgAlaAspAlaProThrValSer 136
Db 304 ACGTTCGGTGGAGGCACCAAGCTGGAATCAACGGGCTGATGCTGCACCAACTGTATCC 363
QY 137 IlePheProSerSerLysLeuGly 145
Db 364 ATCTTCCACCACCTCCAGTAAGCTGGG 390
RESULT 8
US-08-779-784-5
; Sequence 5, Application US/08779784
; Publication No. US20020164325A1
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Moses
; APPLICANT: Miller, David J.
; APPLICANT: Asakura, Kunihiko
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM
; TITLE OF INVENTION: REMYLINATION USING MONOCLONAL AUTOANTIBODIES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,784
; FILING DATE: 07-JAN-1997
```

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/692,084
; FILING DATE: 08-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/236,520
; FILING DATE: 29-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; US-08-779-784-5

Alignment Scores:
Pred. No.: 4,07e-19 Length: 402
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
Gaps: 8

US-08-836-455-2 (1-145) x US-08-779-784-5 (1-402)
QY 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95
Db 232 TCTGGTGTGCCAAAGAGTTTCAGTGGCAGTAGGTCTGGTGCAGATTATTCTCTCACCATC 291
QY 96 SerSerLeuGluSerGluAspPheVal 104
Db 292 AGCAGCCTTGAGTCTGAAGATTGTGA 318

RESULT 9
US-09-924-099-27
; Sequence 27, Application US/09924099
; Patent No. US20020128450A1
; GENERAL INFORMATION:
; APPLICANT: NISHIDA, Yoshihiro
; APPLICANT: OKURA, Takanori
; APPLICANT: TANIMOTO, Tadao
; APPLICANT: KURIMOTO, Masashi
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/924,099
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 27
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(407)
; NAME/KEY: sig peptide
; LOCATION: (1)...(60)

US-08-836-455-2 (1-145) x US-09-924-099-27 (1-407)
QY 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95
Db 226 TCTGGTGTGCCAAAGAGTTTCAGTGGCAGTAGGTCTGGTGCAGATTATTCTCTCACCATC 285
QY 96 SerSerLeuGluSerGluAspPheVal 104
Db 286 AGCAGCCTTGAGTCTGAAGATTGTGA 312

RESULT 10
US-09-990-205-1
; Sequence 1, Application US/09990205
; Patent No. US20020150572A1
; GENERAL INFORMATION:
; APPLICANT: FOON, Kenneth A.
; APPLICANT: CHATTERJEE, Malaya
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF PSORIASIS
; FILE REFERENCE: 304142000501
; CURRENT APPLICATION NUMBER: US/09/990,205
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: U.S. 09/192,838
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: U.S. 60/065,774
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(447)
; NAME/KEY: mat peptide
; LOCATION: (58)...(447)
; US-09-990-205-1

Alignment Scores:
Pred. No.: 4,47e-19 Length: 447
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
Gaps: 10

US-08-836-455-2 (1-145) x US-09-990-205-1 (1-447)
QY 117 ThrPheGlyGlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrValSer 136
Db 361 ACCTTCGGTGGAGGACCAAGCTGGAAATCAACAGGCGTGTGCTGCACCACTGTATCC 420
QY 137 IlePheProProSerSerLysLeuGly 145
Db 421 ATCTCCACCATCCAGTAGCTTGGG 447

RESULT 11
US-10-153-401-1
; Sequence 1, Application US/10153401
; Publication No. US20030114398A1
; GENERAL INFORMATION:
; APPLICANT: Chatterjee, Malaya
; APPLICANT: Foon, Kenneth A.
```

```
; Chatterjee, Sunil K.
; TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE
; TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA
; NUMBER OF SEQUENCE ADDRESSES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/153,401
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/293,533
; FILING DATE: 1999-04-15
; APPLICATION NUMBER: US 08/372,676
; FILING DATE: 1995-01-17
; APPLICATION NUMBER: US 08/591,196
; FILING DATE: 1996-01-16
; ATTORNEY/AGENT INFORMATION:
; NAME: Catherine M. Polizzi
; REGISTRATION NUMBER: 40,130
; REFERENCE/DOCKET NUMBER: 304142000202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..447
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 58
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-153-401-1
Alignment Scores:
Pred. No.: 4,47e-19 Length: 447
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 14 Gaps: 0
US-08-836-455-2 (1-145) x US-10-153-401-1 (1-447)
Qy 117 ThrPheGlyGlyThrLysLeuGluIleLysArgAlaAspAlaProThrValSer 136
Db 361 ACGTTCGGTGAGGACCAAGCTGGAATCAACGGGCTGTGCTGCACCAACTGTATCC 420
Qy 137 IlePheProSerSerLysLeuGly 145
Db 421 ATCTTCCCAACCATCCAGTAGCTGGG 447
RESULT 12
US-09-924-099-19
; Sequence 19, Application US/09924099
; Patent No. US20020128450A1
; GENERAL INFORMATION:
; APPLICANT: NISHIDA, Yoshihiro
; APPLICANT: OKURA, Takanori
; APPLICANT: TANIMOTO, Tadao
; APPLICANT: KURIMOTO, Masashi
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/924,099
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-06-23
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 20
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial DNA to code for the amino acid sequence of
; OTHER INFORMATION: SEQ
; OTHER INFORMATION: ID NO:10
US-09-924-099-20
; GENERAL INFORMATION:
; APPLICANT: NISHIDA, Yoshihiro
; APPLICANT: OKURA, Takanori
; APPLICANT: TANIMOTO, Tadao
; APPLICANT: KURIMOTO, Masashi
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/924,099
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-06-23
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 20
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial DNA to code for the amino acid sequence of
; OTHER INFORMATION: SEQ
; OTHER INFORMATION: ID NO:9
US-09-924-099-19
Alignment Scores:
Pred. No.: 6,77e-19 Length: 711
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.00% Indels: 0
DB: 10 Gaps: 0
US-08-836-455-2 (1-145) x US-09-924-099-19 (1-711)
Qy 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95
Db 556 TCTGGTGTCCCAAGAGGTTTCAGTGGCAGTAGGTCTGGGTGCAGATTATTCTCACCATC 615
Qy 96 SerSerLeuGluSerGluAspPheVal 104
Db 616 AGCAGCCTTGAGTCTGAAGATTGTGA 642
RESULT 13
US-09-924-099-20
; Sequence 20, Application US/09924099
; Patent No. US20020128450A1
; GENERAL INFORMATION:
; APPLICANT: NISHIDA, Yoshihiro
; APPLICANT: OKURA, Takanori
; APPLICANT: TANIMOTO, Tadao
; APPLICANT: KURIMOTO, Masashi
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/924,099
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-06-23
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 20
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial DNA to code for the amino acid sequence of
; OTHER INFORMATION: SEQ
; OTHER INFORMATION: ID NO:10
US-09-924-099-20
```



Alignment Scores:  
Pred. No.: 6.92e-19 Length: 729  
Score: 29.00 Matches: 29  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.00% Indels: 0  
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x US-09-924-099-20 (1-729)

Qy 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95  
Db 556 TCTGGTGTCCCAAGAGTTTCAGTGGCAGTAGGCTGGGTCTCTCAGATTATTTCTCACCATC 615  
Qy 96 SerSerLeuGluSerGluAspPheVal 104  
Db 616 AGCAGCCTTGAGTCTGAAGATTTTGTGA 642

RESULT 14

US-09-887-853-5  
; Sequence 5, Application US/09887853  
; Patent No. US20020168375A1

GENERAL INFORMATION:

APPLICANT: Huston, James S.  
Oppermann, Hermann  
Houston, L. L.  
Ring, David B.

TITLE OF INVENTION: Biosynthetic Binding Proteins For

Imaging

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department

STREET: Exchange Place, 53 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/887.853

FILING DATE: 21-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/133.804

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Kelley, Robin D.

REGISTRATION NUMBER: 34,637

REFERENCE/DOCKET NUMBER: 2054/22

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-248-7477

TELEFAX: 617-248-7100

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 739 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..729

OTHER INFORMATION: /product= "520C9 sfv polypeptide

sequence"

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-887-853-5

Alignment Scores:

Pred. No.: 7.01e-19 Length: 739  
Score: 29.00 Matches: 29  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.00% Indels: 0  
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x US-09-887-853-5 (1-739)

Qy 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95  
Db 565 TCTGGTGTCCCAAGAGTTTCAGTGGCAGTAGGCTGGGTCTGGGTCTCTCAGATTATTTCTCACCATC 624  
Qy 96 SerSerLeuGluSerGluAspPheVal 104  
Db 625 AGTAGCCTTGAGTCTGAAGATTTTGTGA 651

RESULT 15

US-10-268-883-10

; Sequence 10, Application US/10268883

; Publication No. US20030138862A1

GENERAL INFORMATION:

APPLICANT: Tso, J. Yun

APPLICANT: Green, Jennifer Macphate

TITLE OF INVENTION: Anti-HLA-DR Antibodies and the Methods of Using Thereof

FILE REFERENCE: 05882.0062.NPUS01

CURRENT APPLICATION NUMBER: US/10/268,883

CURRENT FILING DATE: 2003-03-26

PRIOR APPLICATION NUMBER: USSN 60/329,178

PRIOR FILING DATE: 2001-10-10

PRIOR APPLICATION NUMBER: USSN 60/331,965

PRIOR FILING DATE: 2001-11-21

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn version 3.1

SEQ ID NO 10

LENGTH: 390

TYPE: DNA

ORGANISM: Mouse

US-10-268-883-10

Alignment Scores:

Pred. No.: 3.67e-18 Length: 390  
Score: 28.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 19.31% Indels: 0  
DB: 12 Gaps: 0

US-08-836-455-2 (1-145) x US-10-268-883-10 (1-390)

Qy 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerAspTyrSerLeuThrIle 95  
Db 232 TCTGGTGTCCCAAGAGTTTCAGTGGCAGTAGGCTGGGTCTGGGTCTCTCAGATTATTTCTCACCATC 291  
Qy 96 SerSerLeuGluSerGluAspPhe 103  
Db 292 AGCAGCCTTGAGTCTGAAGATTTT 315

Search completed: August 31, 2003, 04:04:58

Job time : 193.198 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 30, 2003, 22:20:21 ; Search time 1744.87 Seconds  
(without alignments)  
2019.725 Million cell updates/sec

Title: US-08-836-455-2  
Perfect score: 145  
Sequence: 1 MGAPAIQILGFLLLFPPTGRC.....IKRDAAPTIVSIFPPSSKIG 145

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 1

Total number of hits satisfying chosen parameters: 45562604

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US08836455/runat\_29082003.132953.23055/app\_query.fasta\_1.654  
-DB=EST -OFMT=fastap -SUFFIX=p2noli.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptp  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US08836455.ecgn\_1.1.3596.etrnat\_29082003.132953.23055 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_estc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hun:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gssl:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	30	20.7	594	10	BE309592	BE309592 601094848
2	29	20.0	871	13	BO956722	BO956722 AGENCOURT
3	28	19.3	374	13	BY083003	BY083003 BY083003
C	28	19.3	585	12	BM194777	BM194777 L0700H12-
5	28	19.3	606	13	BO922747	BO922747 AGENCOURT
6	28	19.3	630	10	BF138788	BF138788 601780387
7	28	19.3	707	12	BI250555	BI250555 602993614
8	28	19.3	725	12	BG963055	BG963055 602828068
9	28	19.3	805	12	BI454240	BI454240 603170666
10	28	19.3	849	10	BF583521	BF583521 602101553
11	28	19.3	854	13	BO947692	BO947692 AGENCOURT
12	28	19.3	886	13	BO940987	BO940987 AGENCOURT
13	28	19.3	891	10	BF579422	BF579422 602093833
14	28	19.3	906	10	BF785914	BF785914 602112548
15	28	19.3	908	13	BU524124	BU524124 AGENCOURT
16	28	19.3	926	13	BO959057	BO959057 AGENCOURT
17	28	19.3	935	11	BC031349	BC031349 Mus muscu
18	28	19.3	941	13	BU523453	BU523453 AGENCOURT
19	28	19.3	944	10	BF687485	BF687485 602102475
20	28	19.3	959	13	BO939046	BO939046 AGENCOURT
21	28	19.3	967	10	BF687410	BF687410 602102583
22	28	19.3	975	12	BG961850	BG961850 602826515
23	27	18.6	750	12	BG965050	BG965050 602829112
24	26	17.9	353	13	BY220309	BY220309 BY220309
25	26	17.9	540	12	BG964740	BG964740 602829470
26	26	17.9	553	10	BE309991	BE309991 601091717
27	26	17.9	569	12	BG964957	BG964957 602829274
C	26	17.9	573	12	BM245196	BM245196 K0721C06-
29	26	17.9	574	13	BE305476	BE305476 601099542
C	26	17.9	594	13	BO175706	BO175706 UI-M-DJ2-
31	26	17.9	601	13	BU610021	BU610021 UI-M-DJ2-
C	26	17.9	611	13	BQ175762	BQ175762 UI-M-DJ2-
33	26	17.9	620	10	BE368918	BE368918 601221562
34	26	17.9	641	10	BF134573	BF134573 601784982
35	26	17.9	643	10	BF579280	BF579280 602093454
36	26	17.9	659	10	BF137298	BF137298 601781658
C	26	17.9	671	12	BM194931	BM194931 L0703D04-
38	26	17.9	693	12	BG964281	BG964281 602829076
39	26	17.9	698	12	BH109045	BH109045 602896878
40	26	17.9	706	12	BG965094	BG965094 602829170
41	26	17.9	709	12	BG962745	BG962745 602830268
42	26	17.9	711	10	BF142976	BF142976 601791755
43	26	17.9	712	12	BH100311	BH100311 602858776
44	26	17.9	714	12	BG963548	BG963548 602831226
45	26	17.9	716	12	BG964192	BG964192 602828968

ALIGNMENTS

RESULT 1  
BE309592  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

BE309592 594 bp mRNA linear EST 26-OCT-2000  
601094848F1 NCL\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:3489635 5',  
mRNA sequence.  
BE309592  
BE309592.1 GI:9168025  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 594)

**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>  
 Plate: LLAM8530 row: o column: 12  
 High quality sequence stop: 591.

**FEATURES**

source  
 1. 594  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3489635"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="7 months"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP\_Mam5"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: Sali;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Robin Humphreys,  
 NIH"  
 BASE COUNT 158 a 149 c 143 g 144 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 5.51e-17 Length: 594  
 Score: 30.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 20.69% Indels: 0  
 DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x BE309592 (1-594)

Qy 113 SerProTyrThrPheGlyGlyThrLysLeuGlutLeLysArgAlaAspAlaAla 132  
 |||||||  
 Db 323 AGCTCTCCGTACACGTTCCGGAGGGGGACCAAGCTGGAATAAAACGGCTGCTGCA 382

Qy 133 ProThrValSerIlePheProSerSer 142  
 |||||||  
 Db 383 CCAACTGTATCCATCTTCCACCACCATCCAGT 412

RESULT 2  
 BQ956722  
 LOCUS BQ956722 871 bp mRNA linear EST 21-AUG-2002  
 DEFINITION AGENCOURT 8880951 NCI\_CGAP\_Co24 Mus musculus cDNA clone  
 IMAGE:6396481 5', mRNA sequence.

ACCESSION BQ956722.1 GI:22372200  
 VERSION BQ956722.1  
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 871)  
 TITLE NIH-MGC <http://mgc.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLAM13892 row: n column: 02  
 High quality sequence stop: 630.

**FEATURES**

source  
 Location/Qualifiers  
 1. 871  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6396481"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Co24"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: Sali; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.6 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 218 a 268 c 182 g 193 t 10 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 6.53e-16 Length: 871  
 Score: 29.00 Matches: 29  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 20.00% Indels: 0  
 DB: 13 Gaps: 0

US-08-836-455-2 (1-145) x BQ956722 (1-871)

Qy 114 SerProTyrThrPheGlyGlyThrLysLeuGlutLeLysArgAlaAspAlaAlaPro 133  
 |||||||  
 Db 187 TCCCGGTACACGTTCCGGAGGGGGACCAAGCTGGAATAAAACGGCTGCTGCACCA 246

Qy 134 ThrValSerIlePheProSerSer 142  
 |||||||

Db 247 ACTGTATCCATCTTCCACCACCATCCAGT 273

RESULT 3  
 BQ83003

LOCUS BQ83003 374 bp mRNA linear EST 07-DEC-2002  
 DEFINITION BQ83003 RIKEN full-length enriched, pooled tissues, adult spleen,  
 etc. Mus musculus cDNA clone K630040M6 5', mRNA sequence.

ACCESSION BQ83003  
 VERSION BQ83003.1 GI:26193211  
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 374)

AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
 Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,  
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,  
 Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,  
 Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,  
 Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani,  
 L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,  
 A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,  
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,  
 Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,  
 King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,  
 P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,  
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perle, G.,  
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,  
 Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,  
 B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou,  
 M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale,  
 R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,  
 Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yaginisawa,  
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,  
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura



Sali/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is 1.2 kb. The library was constructed by Yulan Piao (NIA)."

BASE COUNT  
ORIGIN

137 a 121 c 167 g 160 t

#### Alignment Scores:

Pred. No.: 3,49e-15 Length: 585  
Score: 28.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 19.31% Indels: 0  
DB: 12 Gaps: 0

US-08-836-455-2 (1-145) x BM194777 (1-585)

QY 115 ProTyrrPheGlyGlyThrLysLeuGluLeuLysArgAlaAspAlaProThr 134  
DB 571 CCGTACACGTTCCGGAGGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACT 512  
QY 135 ValSerIlePheProSerSer 142  
DB 511 GTATCCATCTCCACCACATCACT 488

#### RESULT 5

BQ922747 606 bp mRNA linear EST 20-AUG-2002  
LOCUS AGENCOURT\_8921909 NCI\_CGAP\_Co24 Mus musculus cDNA clone  
DEFINITION IMAGE:6395978 5', mRNA sequence.

ACCESSION BQ922747 GI:223337778

VERSION BQ922747

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 606)  
NIH-MGC http://mgc.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM13891 row: 1 column: 03  
High quality sequence stop: 605.

#### FEATURES

source

1. 606

Location/Qualifiers

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:6395978"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NCI\_CGAP\_Co24"

/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;

Site\_2: Sali; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.6 kb. Constructed by Life

Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 181 a 169 c 121 g 135 t

#### ORIGIN

Alignment Scores:  
Pred. No.: 3.62e-15 Length: 606  
Score: 28.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

#### Query Match:

DB: 19.31% Indels: 0

13 Gaps: 0

US-08-836-455-2 (1-145) x BQ922747 (1-606)

QY 115 ProTyrrPheGlyGlyThrLysLeuGluLeuLysArgAlaAspAlaProThr 134  
DB 15 CCGTACACGTTCCGGAGGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACT 74

QY 135 ValSerIlePheProSerSer 142

DB 75 GTATCCATCTCCACCACATCACT 98

#### RESULT 6

BQ922747

LOCUS BQ922747

DEFINITION

IMAGE:6395978 5', mRNA sequence.

ACCESSION BQ922747 GI:10977828

VERSION BQ922747

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 630)

NIH-MGC http://mgc.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM9242 row: n column: 21

High quality sequence stop: 628.

Location/Qualifiers

1. 630

source

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="Czech II"

/db\_xref="taxon:10090"

/clone="IMAGE:4008404"

/tissue\_type="tumor, metastatic to mammary"

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP\_Lu30"

/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: NotI;

Site\_2: Sali; transgenic model WNR-1, expression driven by

MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo

dt. Library constructed by Life Technologies.

Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 162 a 158 c 150 g 160 t

ORIGIN

Alignment Scores:

Pred. No.: 3.77e-15 Length: 630

Score: 28.00 Matches: 28

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 19.31% Indels: 0

DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x BQ922747 (1-630)

QY 76 SerGlyValProLysArgPheSerGlySerArgSerGlySerLeuThr 95

DB 222 TCTGGTGTCCCAAGAGTTTCAGTGGCAGTAGGTCTGGGTGAGATTATCTCACCACATC 281

QY 96 SerSerLeuGluSerGluAspPhe 103

```

Db      282  AGCAGCTTGAGTCTGAAGACTTT 305
|||||
RESULT 7
BI250555
LOCUS
DEFINITION
602993614F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5149563 5',
mRNA sequence.
ACCESSION
BI250555
VERSION
BI250555.1 GI:14799016
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 707)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1369 row: k column: 04
High quality sequence stop: 707.
FEATURES
source
1..707
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5149563"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT 198 a 171 c 172 g 166 t
ORIGIN
Alignment Scores:
Pred. No.: 4,24e-15 Length: 707
Score: 28.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.31% Indels: 0
DB: 12 Gaps: 0
US-08-836-455-2 (1-145) x BI250555 (1-707)
QY 115 ProTyrThrPheGlyGlyThrLysLeuGluIleLysArgAlaAspAlaProThr 134
|||||
Db 337 CCTACACGTCGGAGGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACT 396
|||||
QY 135 ValSerIlePheProSerSer 142
|||||
Db 397 GTATCCATCTTCCACCACCATCCAGT 420
|||||
RESULT 8
BG963055
LOCUS
DEFINITION
602828068F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4982825 5',
mRNA sequence.
ACCESSION
BG963055
VERSION
BG963055.1 GI:14350692
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 725)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10986 row: g column: 18
High quality sequence stop: 719.
FEATURES
source
1..725
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4982825"
/lab_host="DHI0B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 195 a 182 c 174 g 174 t
ORIGIN
Alignment Scores:
Pred. No.: 4,35e-15 Length: 725
Score: 28.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.31% Indels: 0
DB: 12 Gaps: 0
US-08-836-455-2 (1-145) x BG963055 (1-725)
QY 115 ProTyrThrPheGlyGlyThrLysLeuGluIleLysArgAlaAspAlaProThr 134
|||||
Db 373 CCGTACACGTCGGAGGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCACCAACT 432
|||||
QY 135 ValSerIlePheProSerSer 142
|||||
Db 433 GTATCCATCTTCCACCACCATCCAGT 456
|||||
RESULT 9
BI454240
LOCUS
DEFINITION
603170666F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5250017 5',
mRNA sequence.
ACCESSION
BI454240
VERSION
BI454240.1 GI:15244896
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 805)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov

```

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM1631 row: d column: 18  
High quality sequence stop: 790.

## FEATURES

source  
1. .805  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5250017"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Mam5"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: Sali;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"  
BASE COUNT 230 a 209 c 186 g 180 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.84e-15 Length: 805  
Score: 28.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 19.31% Indels: 0  
DB: 12 Gaps: 0

US-08-836-455-2 (1-145) x BT454240 (1-805)

Qy 115 ProTyrrPheGlyGlyThrLysLeuGluLeuLysArgAlaAspAlaProThr 134  
|||||  
Db 358 CCGTACACGTTGGAGGGGGACCAAGCTGGAATAAAGGGGTGATGCTGCACCAACT 417

Qy 135 ValSerIlePheProSerSer 142  
|||||  
Db 418 GTATCCATCTCCACCACCATCCAGT 441

RESULT 10

BF583521 849 bp mRNA linear EST 12-DEC-2000  
LOCUS 602101553f1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:4224853 5',  
DEFINITION mRNA sequence.

ACCESSION BF583521.1 GI:11657239

VERSION BF583521

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE NIH-MGC <http://mgc.nci.nih.gov/>

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM9815 row: i column: 14

High quality sequence stop: 627.

## FEATURES

source

1. .849

Location/Qualifiers

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4224853"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NCI\_CGAP\_Co24"

/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;

Site\_2: Sali; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.6 kb. Constructed by Life

Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 274 a 215 c 183 g 177 t

ORIGIN

Alignment Scores:

Pred. No.: 5.11e-15 Length: 849

Score: 28.00 Matches: 28

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 19.31% Indels: 0

DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x BF583521 (1-849)

Qy 115 ProTyrrPheGlyGlyThrLysLeuGluLeuLysArgAlaAspAlaProThr 134

|||||

Db 355 CCGTACACGTTGGAGGGGGACCAACTGGAATAAAGGGGTGATGCTGCACCAACT 414

Qy 135 ValSerIlePheProSerSer 142

|||||

Db 415 GTATCCATCTCCACCACCATCCAGT 438

RESULT 11

BF947692

LOCUS 854 bp mRNA linear EST 21-AUG-2002

DEFINITION AGENCOURT\_8909039 NCI\_CGAP\_Mam2 Mus musculus cDNA clone

IMAGE:6441024 5', mRNA sequence.

ACCESSION BF947692

VERSION BF947692.1 GI:22363170

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>

AUTHORS 1 (bases 1 to 854)

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM13963 row: j column: 01

High quality sequence stop: 671.

Location/Qualifiers

1. .854

source

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N-3"

/db\_xref="taxon:10090"

/clone="IMAGE:6441024"

/tissue\_type="tumor, biopsy sample"

/dev\_stage="5 months"

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP\_Mam2"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: Sali;



Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH

BASE COUNT 234 a 237 c 190 g 193 t  
ORIGIN

## Alignment Scores:

Pred. No.: 5,14e-15 Length: 854  
Score: 28.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 19.31% Indels: 0  
DB: 13 Gaps: 0

US-08-836-455-2 (1-145) x BQ947692 (1-854)

QY 115 ProTyrrThrpheGlyGlyThrLysLeuGlulileysArgAlaaspAlaAProThr 134  
|||||  
Db 361 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAGCGGCTGATGTCGCCCAACT 420

QY 135 ValSerIlePheProSerSer 142  
|||||  
Db 421 GTATCCATCTTCCACCACCATCAGT 444

## RESULT 12

BQ940987

LOCUS

DEFINITION BQ940987 886 bp mRNA linear EST 21-AUG-2002  
IMAGE: 6476176 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 886)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@email.nih.gov

Tissue Procurement: the Cepko Laboratory

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM14016 row: b column: 17

High quality sequence stop: 626.

## FEATURES

source

1. .886  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6476176"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 242 a 243 c 183 g 216 t

ORIGIN

## Alignment Scores:

Pred. No.: 5,33e-15 Length: 886  
Score: 28.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 19.31% Indels: 0

DB: 13 Gaps: 0

US-08-836-455-2 (1-145) x BQ940987 (1-886)

QY 115 ProTyrrThrpheGlyGlyThrLysLeuGlulileysArgAlaaspAlaAProThr 134  
|||||

Db 316 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAGCGGCTGATGTCGCCCAACT 375

QY 135 ValSerIlePheProSerSer 142  
|||||

Db 376 GTATCCATCTTCCACCACCATCAGT 399

## RESULT 13

BF579422

LOCUS

DEFINITION BF579422 891 bp mRNA linear EST 12-DEC-2000  
G02053833F1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:4208144 5',  
mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 891)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@email.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM9772 row: a column: 09

High quality sequence stop: 711.

## FEATURES

source

1. .891  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4208144"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 246 a 226 c 204 g 215 t

ORIGIN

## Alignment Scores:

Pred. No.: 5,36e-15 Length: 891  
Score: 28.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 19.31% Indels: 0  
DB: 10 Gaps: 0

US-08-836-455-2 (1-145) x BF579422 (1-891)

QY 115 ProTyrrThrpheGlyGlyThrLysLeuGlulileysArgAlaaspAlaAProThr 134  
|||||

Db 354 CCGTACACGTTCCGAGGGGGACCAAGCTGGAATAAAGCGGCTGATGTCGCCCAACT 413

QY 135 ValSerIlePheProSerSer 142  
|||||

Db 414 GTATCCATCTTCCACCACCATCAGT 437

```

RESULT 14
BF785914
LOCUS
DEFINITION
602112548F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4240762
5', mRNA sequence.
ACCESSION
BF785914
VERSION
BF785914.1 GI:12090950
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 906)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9855 row: p column: 11
High quality sequence stop: 718.
FEATURES
source
1..906
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4240762"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Kid14"
/notes="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT
243 a 232 c 237 g 194 t
ORIGIN
Alignment Scores:
Pred. No.: 5,46e-15 Length: 906
Score: 28.00 Matches: 28
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.31% Indels: 0
DB: 10 Gaps: 0
US-08-836-455-2 (1-145) x BF785914 (1-906)
QY 115 ProTyrThrPheGlyGlyThrLysLeuGlulileysArgAlaAspAlaProThr 134
|||||
Db 372 CCGTACACGTTCGGAGGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCCCAACT 431
|||||
QY 135 ValSerIlePheProSerSer 142
|||||
Db 432 GTATCCATCTTCCACCATCCAGT 455
|||||
RESULT 15
BU524124
LOCUS
DEFINITION
BU524124
AGENCOURT_10126425 NCI_CGAP_Co24 Mus musculus cDNA clone
IMAGE:6530882 5', mRNA sequence.
ACCESSION
BU524124
VERSION
BU524124.1 GI:22834563
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE
1 (bases 1 to 908)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM14133 row: f column: 02
High quality sequence stop: 604.
FEATURES
source
1..908
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6530882"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT
238 a 270 c 194 g 206 t
ORIGIN
Alignment Scores:
Pred. No.: 5,47e-15 Length: 908
Score: 28.00 Matches: 28
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.31% Indels: 0
DB: 13 Gaps: 0
US-08-836-455-2 (1-145) x BU524124 (1-908)
QY 115 ProTyrThrPheGlyGlyThrLysLeuGlulileysArgAlaAspAlaProThr 134
|||||
Db 173 CCGTACACGTTCGGAGGGGGACCAAGCTGGAATAAAGCGGCTGATGCTGCCCAACT 232
|||||
QY 135 ValSerIlePheProSerSer 142
|||||
Db 233 GTATCCATCTTCCACCATCCAGT 256
|||||
Search completed: August 31, 2003, 02:16:24
Job time : 1747.87 secs

```

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	740	100.0	145	18	AAW27119	Murine monoclonal
2	740	100.0	145	24	AAO16292	Mouse 11D10 anti-
3	727	98.2	145	20	AAW87593	Antibody 11D10 li
4	662	89.5	146	4	AAP30251	Sequence of the le
5	633	85.5	135	21	AAV44599	Mouse light chain
6	527	71.2	129	18	AAE25537	Murine anti-human
7	527	71.2	129	23	AAE20200	Murine 44H104 mab
8	518	70.0	129	23	AAE20204	Murine 44H104 mab
9	508	68.6	246	14	AAAR39569	Sequence of 520C9

```

FT FT /note= "complementarity determining region 2"
FT FT 77..108
FT FT /label= FR3
FT FT /note= "framework region 3"
FT FT 109..117
FT FT /label= CDR3
FT FT /note= "complementarity determining region 3"
FT FT 118..127
FT FT /label= FR1
FT FT /note= "framework region 4"
XX XX
PN WO9722699-A2.
XX XX
PD 26-JUN-1997.
XX XX
PP 19-DEC-1996; 96WO-US20757.
XX XX
PR 20-DEC-1995; 95US-0575762.
PR 26-JAN-1996; 96US-0591965.
PR 13-DEC-1996; 96US-0766350.
XX XX
PA (KENT ) UNIV KENTUCKY.
XX XX
PI Chatterjee M, Chatterjee SK, Foon KA;
XX XX
DR WPI; 1997-341690/31.
DR N-PSDB; AAT85149.
XX XX
PT Monoclonal anti-idiotype antibody 11D10 - elicits immune response
PT against human milk fat globule disease associated tumours,
PT especially breast cancer
XX XX
PS Claim 9; Page 94; 130pp; English.
XX XX
CC This polypeptide sequence comprises the light chain variable region
CC (VL) of monoclonal anti-idiotype antibody 11D10 produced by
CC hybridoma cell line ATCC 12020. 11D10 was obtained by immunising
CC naive mice with MC-10 anti-HMFG antibody to obtain an anti-idiotype
CC response. It elicits an immune response against a specific epitope
CC of a high mol.wt. mucin of human milk fat globule (HMFG). It
CC induces an immunological response to HMFG in mice, rabbits, monkeys
CC and patients with advanced HMFG-associated tumours. Pharmaceutical
CC compositions and vaccines comprising 11D10, 11D10 polypeptides
CC and/or 11D10 polynucleotides (see also AAT85149-50) are claimed.
CC Also claimed are diagnostic kits and methods of using 11D10, 11D10
CC polypeptides and/or 11D10 polynucleotides, including methods of
CC treating HMFG-associated tumours. 11D10 is also used in a claimed
CC method of palliating HMFG-associated disease and in claimed kits to
CC detect or quantify anti-HMFG antibody.
CC (Updated on 25-MAR-2003 to correct PR field.)
XX XX
SQ Sequence 145 AA;
Query Match 100.0%; Score 740; DB 18; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGAPAIQLGFLLLFPCTGCDIQMTQSPSSLSASLGQVSLTCSRASQDIGINLHWLQEP 60
DB 1 MGAPAIQLGFLLLFPCTGCDIQMTQSPSSLSASLGQVSLTCSRASQDIGINLHWLQEP 60
QY 61 DGTIKRLIYATSSLGSGVPRFRFSGRSGSDYSLTISSEDFVAYYCLQYASSPYTFGG 120
DB 61 DGTIKRLIYATSSLGSGVPRFRFSGRSGSDYSLTISSEDFVAYYCLQYASSPYTFGG 120
QY 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145
DB 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145
RESULT 2
AAO16292
ID AAO16292 standard; Protein; 145 AA.
XX XX
SQ Sequence 145 AA;
Query Match 100.0%; Score 740; DB 24; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGAPAIQLGFLLLFPCTGCDIQMTQSPSSLSASLGQVSLTCSRASQDIGINLHWLQEP 60
DB 1 MGAPAIQLGFLLLFPCTGCDIQMTQSPSSLSASLGQVSLTCSRASQDIGINLHWLQEP 60
QY 61 DGTIKRLIYATSSLGSGVPRFRFSGRSGSDYSLTISSEDFVAYYCLQYASSPYTFGG 120
DB 61 DGTIKRLIYATSSLGSGVPRFRFSGRSGSDYSLTISSEDFVAYYCLQYASSPYTFGG 120
QY 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145
DB 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145
RESULT 3
AAO16292
ID AAO16292 standard; Protein; 145 AA.

```

XX AC AAW87593;  
XX DT 16-MAR-1999 (first entry)  
XX XX Antibody 11D10 light chain variable region.  
DE XX Murine; mouse; antibody; light chain; variable region; anti-idiotypic;  
XX KW human milk fat globule; tumour; ovary; lung; pancreas; carcinoma; breast.  
XX KW Mus sp.  
OS XX  
XX FH Key  
FT Peptide 1..20 Location/Qualifiers  
FT Protein 21..999 /note= "signal peptide"  
FT Region 21..43 /note= "mature protein"  
FT Domain 44..54 /label= framework\_1  
FT /label= CDR1  
FT /note= "complementarity determining region 1"  
FT Region 55..69 /label= framework\_2  
FT Misc-difference 55 /note= "encoded by TGG"  
FT Domain 70..76 /label= CDR2  
FT /note= "complementarity determining region 2"  
FT Region 77..108 /label= framework\_3  
FT Domain 109..117 /label= CDR3  
FT /note= "complementarity determining region 3"  
FT Region 118..127 /label= framework\_4  
XX WO9856419-A1.  
PN XX  
XX 17-DEC-1998.  
PD XX  
XX 12-JUN-1998; 98WO-US12250.  
XX 11-JUN-1998; 98US-0096244.  
PR 13-JUN-1997; 97US-0049540.  
XX (KENT ) UNIV KENTUCKY RES FOUND.  
XX Chatterjee M, Foon KA;  
PI WPI; 1999-060029/05.  
XX N-PSDB; AAV83772.  
DR Delaying development of, or treating, HMFg-associated tumours -  
PT using anti-idiotypic antibody 11D10 raised against antibodies to  
PT human milk fat globule protein  
XX  
XX Disclosure; Fig 1; 54pp; English.  
PS  
XX This sequence represents the light chain variable region of the murine  
CC antibody 11D10. This anti-idiotypic antibody is used to delay the  
CC development of, or treat, a human milk fat globule (HMFg) associated  
CC tumour in an individual having low tumour burden. The antibody 11D10  
CC is used to prevent the recurrence of HMFg-associated tumours e.g.  
CC ovarian, non-small cell lung and pancreatic carcinoma, especially for  
CC treating breast tumours.  
XX  
XX Sequence 145 AA;  
SQ  
Query Match 98.2%; Score 727; DB 20; Length 145;  
Best Local Similarity 99.3%; Pred. No. 1.4e-47;  
Matches 144; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGAPAOILGFLLLFFETGTRCDIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQQEP 60  
Db 1 MGAPAOILGFLLLFFETGTRCDIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQQEP 60  
QY 61 DGTIKELIYATSSLSGSGVPRSGSGSDYSLTISLSLEDFVAYCYQYASSPYTFGG 120  
Db 61 DGTIKELIYATSSLSGSGVPRSGSGSDYSLTISLSLEDFVAYCYQYASSPYTFGG 120  
QY 121 GYKLEIKRADAAPTVSIFPPSSKLG 145  
Db 121 GYKLEIKRADAAPTVSIFPPSSKLG 145  
RESULT 4  
AAP30251  
ID AAP30251 standard; peptide; 146 AA.  
XX  
AC AAP30251;  
XX  
DT 25-MAY-1992 (first entry)  
XX  
DE Sequence of the leader, variable region and first 16 AAs of  
DE the constant region of the kappa-chain (light chain) of MOPC41.  
XX  
KW Diagnosis; therapy; immunoglobulin.  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FT Region 1..22 /label= leader  
FT Region 23..130 /label= variable  
FT Region 131..146 /label= constant  
XX EP88994-A.  
PN  
XX 21-SEP-1983.  
PD  
XX 10-MAR-1983; 83EP-0001655.  
PF  
XX 15-MAR-1982; 82US-0358414.  
PR 05-DEC-1983; 83US-0558551.  
XX  
XX (SCHE ) SCHERING CORP.  
XX (DNAX-) DNAX RES INST.  
XX  
XX Moore KW, Zaffaroni A;  
PI WPI; 1983-772290/39.  
XX N-PSDB; AAN30165.  
DR  
XX Transformed expression vectors or plasmid(s) - with double  
PT stranded DNA sequence coding only for desired part of polypeptide  
PT chain  
XX  
XX Example; Page 40-41; 68pp; English.  
PS  
XX The pref. vector or plasmid of the invention has a double-stranded  
CC DNA seq. coding for a variable region of a light or heavy chain of  
CC IgG, or for a variable region of a light or heavy chain of an  
CC immunoglobulin specific for an enzyme or surface protein. The  
CC sequence esp. codes for a variable region of a light chain having 95-  
CC 115 AAs or for a variable region of a heavy chain having 110-125 AAs  
CC esp. including the D region of the heavy chain.  
XX  
XX Sequence 146 AA;  
SQ  
Query Match 89.5%; Score 662; DB 4; Length 146;  
Best Local Similarity 91.6%; Pred. No. 1.1e-42;  
Matches 131; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGAPAIQLGFLLLFPCTRCIDIQMTQSPSSLSASLGORVSLTCSRASODIGINLHWLQOEP 60  
 Db 3 MRAPAIQIFGFLLLFQCTRCIDIQMTQSPSSLSASLGORVSLTCSRASODIGINLHWLQOEP 62  
 QY 61 DGTIKRLIYATSLGSGVPRFSGSRGSDYSLTISSEDEFVAYICLQYASSPYTFGG 120  
 Db 63 DGTIKRLIYATSLGSGVPRFSGSRGSDYSLTISSEDEFVAYICLQYASSPYTFGG 122  
 QY 121 GTKLEIKRADAAPTYSIFPPSSK 143  
 Db 123 GTKLEIKRADAAPTYSIFPPSS 145

RESULT 5  
 AAY44599  
 ID AAY44599 standard; Protein; 135 AA.  
 XX  
 AC AAY44599;  
 DT 04-APR-2000 (first entry)  
 XX  
 DE Mouse light chain variable region encoded by PCR A product.  
 XX  
 KW Mouse light chain variable region; VL; variable region light chain;  
 KW interleukin-18; immunopathies; inflammatory disorder; autoimmune disease;  
 KW mouse; anti-allergic; anti-inflammatory; immunosuppressive;  
 KW hematopoietic; leukocytopenic; antialgic; antipyretic.  
 XX  
 OS Mus musculus.  
 OS Synthetic.  
 XX  
 PN EP974600-A2.  
 XX  
 PD 26-JAN-2000.  
 XX  
 XX 24-JUN-1999; 99EP-0304977.  
 XX  
 XX 24-JUN-1998; 98JP-0177580.  
 PR 12-OCT-1998; 98JP-0289044.  
 PR 22-DEC-1998; 98JP-0365023.  
 XX  
 PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
 XX  
 PI Nishida Y, Okura T, Tanimoto T, Kurimoto M;  
 XX  
 DR WPI; 2000-118341/11.  
 DR N-PSDB; AAZ49548.  
 XX  
 XX New artificially produced peptide for neutralizing biological activity  
 PT of interleukin-18, useful for treating and preventing immunopathies,  
 PT inflammatory disorders and autoimmune diseases -  
 XX  
 PS Disclosure; Page 28-29; 32pp; English.  
 XX  
 CC The present sequence is the mouse light chain variable region. This was  
 CC encoded by a recombinant DNA is derived from PCR A which amplifies  
 CC antibody light chain variable region (VL). The transformant produced  
 CC using the VL gene was used to transform competent E. coli cells. The peptide  
 CC produced by transformants neutralises interleukin-18. This is useful for  
 CC treating and preventing immunopathies, inflammatory disorders and  
 CC autoimmune diseases which are caused by excessive immunoreaction. The  
 CC peptide has anti-allergic, anti-inflammatory, immunosuppressive,  
 CC hematopoietic, leukocytopenic, antialgic, antipyretic and hepatic  
 CC -function improving activities.  
 XX  
 SQ Sequence 135 AA;

Query Match 85.58; Score 633; DB 21; Length 135;  
 Best Local Similarity 92.6%; Pred. No. 1.6e-40;  
 Matches 125; Conservative 2; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 MGAPAIQLGFLLLFPCTRCIDIQMTQSPSSLSASLGORVSLTCSRASODIGINLHWLQOEP 60  
 Db 1

Db 1 MRAPAIQIFGFLLLFPCTRCIDIQMTQSPSSLSASLGORVSLTCSRASODIGINLHWLQOEP 60  
 QY 61 DGTIKRLIYATSLGSGVPRFSGSRGSDYSLTISSEDEFVAYICLQYASSPYTFGG 120  
 Db 61 DGTIKRLIYATSLGSGVPRFSGSRGSDYSLTISSEDEFVAYICLQYASSPYTFGG 120  
 QY 121 GTKLEIKRADAAPT 135  
 Db 121 GTKLEIKRADAAPT 135

RESULT 6  
 AAW22537  
 ID AAW22537 standard; Protein; 129 AA.  
 XX  
 AC AAW22537;  
 DT 03-NOV-1997 (first entry)  
 XX  
 DE Murine anti-human class II monoclonal antibody 44H104 VL chain.  
 XX  
 KW Antibody; light chain; variable region; hybridoma cell line 44H104;  
 KW immune response; enhance; stimulate; vaccine; immunodiagnosis;  
 KW antigen delivery.  
 XX  
 OS Mus musculus.  
 XX WO9640941-A1.  
 PN 19-DEC-1996.  
 PD 07-JUN-1996; 96WO-CA00400.  
 PF 07-JUN-1995; 95US-0483576.  
 PR (CONN-) CONNAUGHT LAB LTD.  
 XX  
 PA Anand NN, Barber BH, Caterini JE, Cates GC, Klein MH;  
 PI WPI; 1997-077271/07.  
 DR N-PSDB; AAT77851.  
 DR  
 XX Recombinant conjugate antibody mol., modified for delivering an  
 PT antigen - elicits enhanced immune response without the use of  
 PT adjuvant to generate antibodies which are useful in vaccines or  
 PT immuno:diagnosis  
 PS Example 1; Fig 1A; 64pp; English.  
 XX  
 CC Novel recombinant conjugate antibody molecules comprise a monoclonal  
 CC antibody specific for a surface structure of antigen presenting  
 CC cells (APC), genetically modified to contain at least one antigen  
 CC exclusively at one or more preselected sites. The conjugate is capable  
 CC of delivering the antigen to APC and eliciting an immune response to  
 CC the antigen. The new conjugates are useful as vaccines and are able  
 CC to elicit an enhanced immune response without the use of an adjuvant.  
 CC In a specific example, a conjugate was constructed using the murine  
 CC anti-human class II monoclonal antibody secreted by hybridoma  
 CC 44H104. The peptide CLTB36 was chosen as antigen; it consists of  
 CC a tandemly linked T and B cell epitope derived from HIV MN strain.  
 CC The present sequence represents the light chain variable region  
 CC from 44H104 which was used in the preparation of a conjugate with  
 CC antigen CLTB36.  
 XX  
 SQ Sequence 129 AA;

Query Match 71.2%; Score 527; DB 18; Length 129;  
 Best Local Similarity 81.1%; Pred. No. 1.5e-32;  
 Matches 103; Conservative 7; Mismatches 17; Indels 0; Gaps 0;  
 QY 1 MGAPAIQLGFLLLFPCTRCIDIQMTQSPSSLSASLGORVSLTCSRASODIGINLHWLQOEP 60  
 Db 3 MRYPAHVFGFLLLWFFCTRCIDIQMTQSPSSLSASLGORVSLTCSRASODIGINLHWLQOEP 62



Db 3 MRVPAHVFGELLWFPGTRCDIQMTOSPSLSASLGQVSLTCRASQEIISGYLTWLOQKP 62  
 QY 61 DGTIKRLIYATSLGSGVPRFRSGSDYSLTISSESEDFVAYCYLOYASSPYTFGG 120  
 Db 63 DGTIKRLVYAATLDGSGVPRFRSGSDYSLTISSESEDFADYCYLOYTNYPLTFGA 122  
 QY 121 GTKLEIK 127  
 Db 123 GTKLELK 129  
 RESULT 9  
 AAR39569  
 ID AAR39569 standard; Protein; 246 AA.  
 XX  
 AC AAR39569;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 07-FEB-1994 (first entry)  
 XX  
 DE Sequence of 520C9 sFv protein.  
 XX  
 KW Single chain Fv polypeptide; VH-VL heterodimer; immunoglobulin; Ig;  
 KW biosynthetic single polypeptide chain binding site; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN W09316185-A2.  
 XX  
 PD 19-AUG-1993.  
 XX  
 PF 05-FEB-1993; 93WO-US01055.  
 XX  
 PR 06-FEB-1992; 92US-0831967.  
 XX  
 PA (CETU ) CETUS ONCOLOGY CORP.  
 PA (CREA-) CREATIVE BIOMOLECULES INC.  
 XX  
 PI Houston LL, Huston JS, Oppermann H, Ring DB;  
 XX  
 DR WPI; 1993-272889/34.  
 DR N-PSDB; AAQ46084.  
 XX  
 PT New single chain Fv polypeptide binding to C-erbB-2 tumour  
 PT antigen - for imaging or treating breast or ovarian cancer etc.  
 XX  
 PS Claim 4; pages 60-61; 87pp; English.  
 XX  
 CC C-erbB-2 refers to a protein antigen expressed on the surface of  
 CC tumour cells, such as breast and ovarian tumour cells, which is an  
 CC approx. 200,000 mol. wt. acidic glycoprotein having an isoelectric  
 CC pt. of about 5.3 (see AAQ46083, AAR39568). A single chain Fv (sFv)  
 CC is a covalently linked VH-VL heterodimer which is expressed from  
 CC a gene fusion including VH- and VL- encoding genes connected by  
 CC a peptide-encoding linker. Such linker sequences are set forth in  
 CC AA residues 116-135 in AAR39569, which includes part of the 16 AA  
 CC linker sequences in AAR39572. Using AAQ46084 for the 520C9 monoclonal  
 CC antibody, a single chain polypeptide can be produced having a  
 CC binding affinity for a C-erbB-2 related antigen. 'X' in AAR39569  
 CC refers to the location of a stop codon in AAQ46084.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 246 AA;  
 Query Match 68.6%; Score 508; DB 14; Length 246;  
 Best Local Similarity 90.2%; Pred. No. 7.6e-31;  
 Matches 101; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
 QY 21 DIQMTOSPSLSASLGQVSLTCRASQDIGINHLWLOQEPDGTIKRLIYATSLGSGVPR 80  
 Db 134 DIQMTOSPSLSASLGQVSLTCRASQDIGINHLWLOQEPDGTIKRLIYATSLDGVPR 193  
 QY 81 RFGSGRSGSDYSLTISSESEDFVAYCYLOYASSPYTFGGTLEIKRADAA 132

Db 194 RFGSGRSGSDYSLTISSESEDFVAYCYLOYAIFPYTFGGTLEIKRADXS 245  
 RESULT 10  
 AAW02280  
 ID AAW02280 standard; Protein; 243 AA.  
 XX  
 AC AAW02280;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 29-OCT-1996 (first entry)  
 XX  
 DE 520C9 anti-c-erbB-2 two single chain Fv construct.  
 XX  
 KW 520C9; anti-c-erbB-2 monoclonal antibody; single chain Fv; sFv;  
 KW construct; polypeptide linker; C-terminal amino acid sequence;  
 KW in vivo imaging; drug targeting experiment; homodimer;  
 KW increased; binding avidity; tissue retention time.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 118..133  
 FT /label= linker  
 XX  
 PN US5534254-A.  
 XX  
 PD 09-JUL-1996.  
 XX  
 PF 07-OCT-1993; 93US-0133804.  
 XX  
 PR 07-OCT-1993; 93US-0133804.  
 PR 06-FEB-1992; 92US-0831967.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PA (CREA-) CREATIVE BIOMOLECULES INC.  
 XX  
 PI Houston LL, Huston JS, Oppermann H, Ring DB;  
 XX  
 DR WPI; 1996-333194/33.  
 DR N-PSDB; AAT36880.  
 XX  
 PT Compens. contg. antigen-targeting antibody fragment constructs -  
 PT comprising dimer of single-chain Fv fragments  
 XX  
 PS Example 1; Columns 33-36; 30pp; English.  
 XX  
 CC Variable heavy (VH) and variable light (VL) genes were cloned from  
 CC a 520C9 hybridoma cDNA library, using probes directed toward the  
 CC antibody constant and joining regions. A two single chain Fv (sFv)  
 CC gene was constructed by connecting the VH and VL genes with a  
 CC Ser rich polypeptide linker. The resulting 520C9 two sFv gene,  
 CC which encodes the present sequence, was inserted into an expression  
 CC vector, transformed into E. coli, and protein expression induced by  
 CC the addn. of IPTG to the culture medium.  
 CC A compsn. comprising a carrier and the 2 sFv protein prod. can be  
 CC used for in vivo imaging, and drug targeting experiments. The  
 CC 2 sFv protein prod. is a homodimer, in which both fragments target  
 CC the same antigen, therefore giving greater binding avidity and  
 CC longer tissue retention times, compared to individual sFv protein  
 CC prod. fragments.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 243 AA;  
 Query Match 68.5%; Score 507; DB 17; Length 243;  
 Best Local Similarity 91.8%; Pred. No. 8.9e-31;  
 Matches 101; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
 QY 21 DIQMTOSPSLSASLGQVSLTCRASQDIGINHLWLOQEPDGTIKRLIYATSLGSGVPR 80  
 Db 134 DIQMTOSPSLSASLGQVSLTCRASQDIGINHLWLOQEPDGTIKRLIYATSLDGVPR 193



QY 81 RFGSGRSGDYSLTISSEDFVAYICLQYASSPYTFGGTKLEIKRAD 130  
 |||||  
 Db 194 RFGSGRSGDYSLTISSEDFVAYICLQYAIFFYTFGGTNLEIKRAD 243  
 |||||

RESULT 11  
 AAW53170  
 ID AAW53170 standard; Protein; 243 AA.  
 XX  
 AC AAW53170;  
 XX  
 DT 16-JUL-1998 (first entry)  
 XX  
 DE 520C9 anti-c-erbB-2 sfv' dimeric construct protein sequence.  
 KW Antigen imaging; single chain Fv; sfv; linker; dimeric; cancer;  
 KW c-erbB-2; tumour; diagnosis; ss.  
 XX  
 OS Synthetic.  
 OS Mus sp.  
 PN US5753204-A.  
 XX  
 PD 19-MAY-1998.  
 XX  
 PF 05-JUN-1995; 95US-0461838.  
 PR 07-OCT-1993; 93US-0133804.  
 PR 06-FEB-1992; 92US-0831967.  
 PR 05-JUN-1995; 95US-0461838.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PA (CREA-) CREATIVE BIOMOLECULES INC.  
 XX  
 PI Houston LL, Huston JS, Oppermann H, Ring DB;  
 XX  
 DR WPI; 1998-311318/27.  
 DR N-PSDB; AAV21798.  
 XX  
 PT Imaging of antigens in vivo - using dimers of single-chain antibody  
 PT Fv fragments  
 XX  
 PS Example 1; Columns 33-36; 30pp; English.  
 CC This represents the protein sequence of a 520C9 sfv' (single chain Fv)  
 CC construct. This was constructed by connecting the Vh and Vl genes with  
 CC a DNA sequence encoding a Serine rich linker. 520C9 is a monoclonal  
 CC antibody useful in targeting c-erbB-2 antigen. This dimeric construct  
 CC can be used in the methods of invention of imaging a preselected antigen  
 CC expressed in a mammal. The methods are used in magnetic resonance imaging  
 CC of c-erbB-2 or related antigens in cancer diagnosis. The biosynthetic  
 CC constructs have enhanced properties as in vivo targeting agents in  
 CC comparison with intact monoclonal antibodies or their Fab fragments. The  
 CC dimeric constructs permit the in vivo targeting of an epitope on an  
 CC antigen with greater apparent avidity, including greater tumour  
 CC specificity, tumour localisation and tumour retention properties than  
 CC that of the Fab fragment having the same CDRs as the construct.  
 XX  
 SQ Sequence 243 AA;  
 Query Match 68.5%; Score 507; DB 19; Length 243;  
 Best Local Similarity 91.8%; Pred. No. 8.9e-31;  
 Matches 101; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 21 DIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQEPDGTIKRLIYATSSLSGVPK 80  
 |||||  
 Db 134 DIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQEPDGTIKRLIYATSSLSGVPK 193  
 |||||

QY 81 RFGSGRSGDYSLTISSEDFVAYICLQYASSPYTFGGTKLEIKRAD 130  
 |||||  
 Db 194 RFGSGRSGDYSLTISSEDFVAYICLQYAIFFYTFGGTNLEIKRAD 243  
 |||||

RESULT 13  
 AAY44587  
 ID AAY44587 standard; Protein; 108 AA.  
 XX  
 AC AAY44587;  
 XX  
 DT 04-APR-2000 (first entry)  
 XX  
 DE Mouse anti-IL-18 antibody light chain variable region.

RESULT 12  
 AAW80424  
 ID AAW80424 standard; Protein; 243 AA.  
 XX  
 AC AAW80424;  
 XX  
 DT 28-JAN-1999 (first entry)  
 XX  
 DE 520C9 sfv sequence.  
 XX  
 KW 520C9 sfv; antigen; tumour cell; antibody 520C9;  
 KW targeted delivery; antigen-expressing cell.  
 XX  
 OS Synthetic.  
 XX  
 PN US5837846-A.  
 XX  
 PD 17-NOV-1998.  
 XX  
 PF 05-JUN-1995; 95US-0461386.  
 XX  
 PR 07-OCT-1993; 93US-0133804.  
 PR 06-FEB-1992; 92US-0831967.  
 PR 05-JUN-1995; 95US-0461386.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PA (CREA-) CREATIVE BIOMOLECULES INC.  
 XX  
 PI Houston LL, Huston JS, Oppermann H, Ring DB;  
 XX  
 DR WPI; 1999-023541/02.  
 DR N-PSDB; AAV63399.  
 XX  
 PT Nucleic acid encoding single-chain Fv fragment specific for antigens  
 PT - and having C-terminal tail for crosslinking to form dimer with  
 PT improved pharmacokinetic properties, used to deliver drugs and  
 PT imaging agents, especially to tumours  
 XX  
 PS Example 1; Columns 35-36; 29pp; English.  
 CC The present sequence represents an antibody 520C9 sfv. Variable heavy  
 CC and light sequences of antibody 7520C9 are connected, together with a  
 CC serine linker, to produce the present single chain Fv gene. The present  
 CC sequence exemplifies the invention. Dimers of the single chain Fv are  
 CC used for targeted delivery of drugs or imaging agents (e.g. cytotoxins,  
 CC prodrugs or 99m-technetium) to antigen-expressing cells, particularly  
 CC for treatment or diagnosis of tumours (especially of ovary or breast).  
 XX  
 SQ Sequence 243 AA;  
 Query Match 68.5%; Score 507; DB 20; Length 243;  
 Best Local Similarity 91.8%; Pred. No. 8.9e-31;  
 Matches 101; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 21 DIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQEPDGTIKRLIYATSSLSGVPK 80  
 |||||  
 Db 134 DIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQEPDGTIKRLIYATSSLSGVPK 193  
 |||||

QY 81 RFGSGRSGDYSLTISSEDFVAYICLQYASSPYTFGGTKLEIKRAD 130  
 |||||  
 Db 194 RFGSGRSGDYSLTISSEDFVAYICLQYAIFFYTFGGTNLEIKRAD 243  
 |||||

RESULT 13  
 AAY44587  
 ID AAY44587 standard; Protein; 108 AA.  
 XX  
 AC AAY44587;  
 XX  
 DT 04-APR-2000 (first entry)  
 XX  
 DE Mouse anti-IL-18 antibody light chain variable region.

XX Anti-Interleukin-18 antibody; IL-18; light chain variable region; VL;  
 KW Hybridoma #125-2H; mouse; monoclonal antibody #125-2HmAb; interleukin-18;  
 KW antinflammatory; immunosuppressive; leucocytopenic; antiallgic;  
 KW antipyretic; antiallergic; hepatotropic; treatment; autoimmune disease;  
 KW immunopathy; inflammatory disorder; immunoreaction.  
 OS Mus musculus.  
 XX  
 XX EP974600-A2.  
 XX  
 XX 26-JAN-2000.  
 XX  
 XX 24-JUN-1999; 99EP-0304977.  
 XX  
 XX 24-JUN-1998; 98JP-0177580.  
 PR 12-OCT-1998; 98JP-0289044.  
 PR 22-DEC-1998; 98JP-0365023.  
 XX  
 PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
 XX  
 XX Nishida Y, Okura T, Tanimoto T, Kurimoto M;  
 XX  
 XX WPI; 2000-118341/11.  
 DR N-PSDB; AAZ49534.  
 XX  
 XX New artificially produced peptide for neutralizing biological activity  
 of interleukin-18, useful for treating and preventing immunopathies,  
 PT inflammatory disorders and autoimmune diseases -  
 PT  
 XX Claim 5; Page 17; 36pp; English.  
 XX  
 XX The present protein sequence is the mouse anti-Interleukin-18 antibody  
 CC light chain variable region (VL) encoded by cDNA derived from hybridoma  
 CC #125-2H. The nucleotide sequence is used in the production of recombinant  
 CC monoclonal antibody #125-2HmAb, which is capable of neutralising  
 CC biological activities of interleukin-18. The antibody has antiallgic,  
 CC antinflammatory, immunosuppressive, leucocytopenic, antipyretic,  
 CC antiallergic and hepatotropic activity and can be used for prevention and  
 CC treatment of autoimmune diseases, immunopathies and inflammatory  
 CC disorders caused by excessive immunoreaction.  
 XX  
 XX Sequence 108 AA;  
 Query Match 68.4%; Score 506; DB 21; Length 108;  
 Best Local Similarity 92.6%; Pred. No. 4.8e-31;  
 Matches 100; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 QY 21 DIQWTSPLSASLGORVSLTCRASODIGINLHWLQOEPDGTIKRLIYATSSLGSPVK 80  
 Db 1 DIQWTSPLSASLGORVSLTCRASODIGSKLYLWLOQEPDGTIKRLIYATSSLDGSPVK 60  
 QY 81 RFSGRSGSDYSLTISSEDFVAYCYQYASSPYTFGGGKLEIKR 128  
 Db 61 RFSGRSGSDYSLTISSEDFVYCYQYASSPYTFGGGKLEIKR 108  
 RESULT 14  
 AAU72863  
 ID AAU72863 standard; Protein; 252 AA.  
 XX  
 XX AAU72863;  
 XX  
 XX 26-FEB-2002 (first entry)  
 XX  
 XX P4-14 single chain Fv.  
 XX  
 XX Human; NKG2D receptor complex; cancer; infectious disease; tumour;  
 KW autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;  
 KW intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;  
 KW prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;  
 KW sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; DAP10;  
 KW helminth; cytostatic; antimicrobial; immunomodulatory; 11B2D10; 6H7E7;

KW 8G7C10; 6E5A7; 11B2D10x4-7; 8G7C10x4-7; 6E5A7x4-7; P4-2; P4-3; P4-14;  
 KW P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10xP4-3; 3B10xP4-14;  
 KW p53 tetramerisation domain; 3B10xP5-2; 3B10xP5-23.  
 XX Homo sapiens.  
 XX  
 XX WO200171005-A2.  
 XX  
 XX 27-SEP-2001.  
 XX  
 XX 26-MAR-2001; 2001WO-EP03414.  
 XX  
 XX 24-MAR-2000; 2000EP-0106467.  
 PR  
 XX (KUFE/) KUFER P.  
 PA  
 XX Kufer P, Riettmueller G, Lutterbuese R, Borschert K, Kischel R;  
 PI Mayer M, Hofmeister R;  
 PI  
 XX WPI; 2002-055119/07.  
 DR N-PSDB; AAS97137.  
 XX  
 XX Multifunctional polypeptides comprising binding sites that specifically  
 PT recognise extracellular groups of the NKG2D receptor complex and  
 PT domains which function as receptors or ligands, useful for treating  
 PT cancers and infectious diseases -  
 PT  
 XX Example 7; Fig 16; 114pp; English.  
 XX  
 XX The invention relates to a multifunctional polypeptide comprising a  
 CC domain with a binding site that specifically recognises an extracellular  
 CC group of the NKG2D receptor complex and a second domain which functions  
 CC as a receptor or ligand. The polypeptide and its associated  
 CC polynucleotide are used for the preparation of a pharmaceutical  
 CC composition for the treatment of cancer, infections and/or autoimmune  
 CC conditions. The cancer may be a tumour of the head and neck, stomach,  
 CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,  
 CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,  
 CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.  
 CC The infectious diseases can be caused by viruses, bacteria, fungi,  
 CC protozoa or helminths. The autoimmune diseases include multiple  
 CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior  
 CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent  
 CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and  
 CC autoimmune hepatitis. Sequences AAU72820-AAU72875 represent the NKG2D  
 CC receptor and the polypeptides of the invention.  
 XX  
 XX Sequence 252 AA;  
 Query Match 68.2%; Score 505; DB 23; Length 252;  
 Best Local Similarity 86.2%; Pred. No. 1.3e-30;  
 Matches 100; Conservative 7; Mismatches 9; Indels 0; Gaps 0;  
 QY 17 GTRCDIQMTQSPSSLSASLGORVSLTCRASODIGINLHWLQOEPDGTIKRLIYATSSLSGS 76  
 Db 130 GGSELVMTQSPSSLSASLGORVSLTCRASODIGSSLNWLOQEPDGTIKRLIYATSSLSGS 189  
 QY 77 GVPKRFSGSRGSDYSLTISSEDFVAYCYQYASSPYTFGGGKLEIKRADA 132  
 Db 190 GVPKRFSGSRGSDYSLTISSEDFVDYCYQYASSPYTFGGGKLEIKRTSS 245  
 RESULT 15  
 AAU72872  
 ID AAU72872 standard; Protein; 499 AA.  
 XX  
 XX AAU72872;  
 XX  
 XX 26-FEB-2002 (first entry)  
 XX  
 XX 3B10xP4-14 bispecific single chain Fv.  
 DE  
 XX Human; NKG2D receptor complex; cancer; infectious disease; tumour;

KW autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;  
KW intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;  
KW prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; FV;  
KW sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; DAP10;  
KW helminth; cytostatic; antimicrobial; immunomodulatory; 11B2D10; 6H7E7;  
KW 8G7C10; 6E5A7; 11B2D10x4-7; 8G7C10x4-7; 8E5A7x4-7; P4-2; P4-3; P4-14;  
KW P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10XP4-3; 3B10XP4-14;  
KW P53 tetramerisation domain; 3B10XP5-2; 3B10XP5-23.  
XX Homo sapiens.  
OS  
XX  
XX WO200171005-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 26-MAR-2001; 2001WO-EP03414.  
PF  
XX  
XX 24-MAR-2000; 2000EP-0106467.  
PR  
XX  
XX (KUFE/) KUFER P.  
PA  
XX  
XX Kufer P, Riethmueller G, Lutterbuese R, Borschert K, Kischel R;  
PI Mayer M, Hofmeister R;  
PI  
XX  
XX WPI; 2002-055119/07.  
DR N-PSDB; AAS97146.  
DR  
XX  
XX Multifunctional polypeptides comprising binding sites that specifically  
PT recognise extracellular groups of the NKG2D receptor complex and  
PT domains which function as receptors or ligands, useful for treating  
PT cancers and infectious diseases -  
XX  
XX Example 5; Fig 16; 114pp; English.  
XX  
XX The invention relates to a multifunctional polypeptide comprising a  
CC domain with a binding site that specifically recognises an extracellular  
CC group of the NKG2D receptor complex and a second domain which functions  
CC as a receptor or ligand. The polypeptide and its associated  
CC polynucleotide are used for the preparation of a pharmaceutical  
CC composition for the treatment of cancer, infections and/or autoimmune  
CC conditions. The cancer may be a tumour of the head and neck, stomach,  
CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,  
CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,  
CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.  
CC The infectious diseases can be caused by viruses, bacteria, fungi,  
CC protozoa or helminths. The autoimmune diseases include multiple  
CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior  
CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent  
CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and  
CC autoimmune hepatitis. Sequences AAU72820-AAU72875 represent the NKG2D  
CC receptor and the polypeptides of the invention.  
XX  
SQ Sequence 499 AA;

Query Match 68.2%; Score 505; DB 23; Length 499;  
Best Local Similarity 86.2%; Pred. No. 2.5e-30;  
Matches 100; Conservative 7; Mismatches 9; Indels 0; Gaps 0;  
QY 17 GTRCDIOMTQSPSSLSASIGQRVSLTCRASQDIGINLHWLQQEPDGTIKRLIYATSSLGS 76  
Db 377 GGSSELVMTQSPSSLSASIGQRVSLTCRASQDIGINLHWLQQEPDGTIKRLIYATSSLGS 436  
QY 77 GVPKRFSGSRGSDYSLTISSESEDFVAYCYQYASSPYTFGGGTKEIKRADA 132  
Db 437 GVPKRFSGSRGSDYSLTISSESEDFVYCYQYASSPYTFGGGTKEIKRTSS 492

Search completed: August 30, 2003, 22:04:04  
Job time : 36.547 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 21:57:15 ; Search time 23.3557 Seconds  
(without alignments)  
849.188 Million cell updates/sec

Title: US-08-836-455-2  
Perfect score: 740  
Sequence: 1 MGAPAIQLGFLLLFPCTRC.....IKRADAPTIVIPPPSKLG 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 510680 seqs, 136781880 residues

Total number of hits satisfying chosen parameters: 510680

Minimum DB seq length: 0

Maximum Match 100%

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Published: Applications\_AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	740	100.0	145	9	US-09-861-294-2
2	740	100.0	145	12	US-10-367-506-2
3	601	81.2	130	8	US-08-779-784-35
4	541	73.1	130	12	US-10-268-883-11
5	518	70.0	129	9	US-09-007-093-2
6	507	68.5	243	10	US-09-887-853-6
7	506	68.4	108	10	US-09-924-099-1
8	502	67.8	237	10	US-09-924-099-9
9	502	67.8	243	10	US-09-924-099-10
10	496	67.0	109	10	US-09-943-906-74
11	473	63.9	236	10	US-09-859-053-30
12	471	63.6	112	12	US-10-355-780-1
13	471	63.6	142	9	US-09-797-481-2
14	471	63.6	142	9	US-09-844-736-4
15	471	63.6	142	15	US-10-162-396-4

16	464	62.7	108	12	US-10-268-883-12	Sequence 12, Appl
17	461	62.3	131	8	US-08-779-784-21	Sequence 21, Appl
18	455.5	61.6	241	15	US-10-221-945-1	Sequence 1, Appl
19	445	60.1	234	9	US-09-740-002-24	Sequence 24, Appl
20	444	60.0	276	9	US-09-766-543-12	Sequence 12, Appl
21	442	59.7	109	10	US-09-943-906-73	Sequence 73, Appl
22	439	59.3	127	12	US-10-268-883-5	Sequence 5, Appl
23	438	59.2	236	15	US-10-006-593-69	Sequence 69, Appl
24	436	58.9	128	9	US-09-764-304-10	Sequence 10, Appl
25	436	58.9	128	9	US-09-764-304-19	Sequence 19, Appl
26	436	58.9	128	15	US-10-265-713-10	Sequence 10, Appl
27	436	58.9	128	15	US-10-265-713-19	Sequence 19, Appl
28	433	58.5	127	15	US-10-084-139-2	Sequence 2, Appl
29	423.5	57.2	234	9	US-09-740-002-26	Sequence 26, Appl
30	422	57.0	214	15	US-10-153-382-19	Sequence 19, Appl
31	421	56.9	237	9	US-09-056-160B-100	Sequence 100, App
32	421	56.9	491	14	US-10-011-125-2	Sequence 2, Appl
33	419	56.6	237	15	US-10-020-786-8	Sequence 8, Appl
34	419	56.6	237	15	US-10-227-694-4	Sequence 4, Appl
35	417	56.4	214	9	US-09-754-998-1	Sequence 1, Appl
36	416	56.2	237	15	US-10-020-786-10	Sequence 10, Appl
37	415	56.1	121	10	US-09-797-941A-4	Sequence 4, Appl
38	409	55.3	237	9	US-09-940-166A-6	Sequence 6, Appl
39	409	55.3	237	15	US-10-227-694-1	Sequence 1, Appl
40	407	55.0	214	9	US-09-940-166A-2	Sequence 2, Appl
41	407	55.0	214	9	US-09-811-384-11	Sequence 11, Appl
42	407	55.0	214	12	US-10-356-974-1	Sequence 1, Appl
43	407	55.0	214	15	US-10-253-366-1	Sequence 1, Appl
44	407	55.0	214	15	US-10-316-694-1	Sequence 1, Appl
45	406	54.9	107	15	US-10-056-794-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1

US-09-861-294-2  
; Sequence 2, Application US/09861294  
; Patent No. US20020098190A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TITLE OF INVENTION: TUMORS BEARING HMFPG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/09/861,294  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 145  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(20)  
US-09-861-294-2

Query Match	100.0%;	Score 740;	DB 9;	Length 145;
Best Local Similarity	100.0%;	Pred. No. 8.9e-59;		
Matches 145;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	MGAPAIQLGFLLLFPCTRCDIOMTQSPSSLSASLGORVSLTCRASODIGINLHLWQEP	60	
Db	1	MGAPAIQLGFLLLFPCTRCDIOMTQSPSSLSASLGORVSLTCRASODIGINLHLWQEP	60	
Oy	61	DGTRKLIYATSSLSGSGVPRKFRSGSRGSDYSLTISSEDFVAYCYCLQYASSPYTFGG	120	
Db	61	DGTRKLIYATSSLSGSGVPRKFRSGSRGSDYSLTISSEDFVAYCYCLQYASSPYTFGG	120	

Qy 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145  
Db 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145

## RESULT 2

US-10-367-506-2  
; Sequence 2, Application US/10367506  
; Publication No. US20030152575A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; FILE OF INVENTION: TUMORS BEARING HMFG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/10/367,506  
; CURRENT FILING DATE: 2003-02-13  
; PRIOR APPLICATION NUMBER: US/09/861,294  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ. ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 145  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(20)  
US-10-367-506-2

Query Match 100.0%; Score 740; DB 12; Length 145;  
Best Local Similarity 100.0%; Pred. No. 8.9e-59;  
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MGAPAIQLGFLLLFPCTRCDIQMTQSPSSLSASISGQVSVLTCSRASODIGINLHWLOEP 60  
Db 1 MGAPAIQLGFLLLFPCTRCDIQMTQSPSSLSASISGQVSVLTCSRASODIGINLHWLOEP 60  
Qy 61 DGTIKRLIYATSSILGSGVPRKFRSGSDYSLTISSEDFVAYYCLQYASSPYTFGG 120  
Db 61 DGTIKRLIYATSSILGSGVPRKFRSGSDYSLTISSEDFVAYYCLQYASSPYTFGG 120  
Qy 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145  
Db 121 GTKLEIKRADAAPTVSIFFPPSSKLG 145

## RESULT 3

US-08-779-784-35  
; Sequence 35, Application US/08779784  
; Publication No. US20020164325A1  
; GENERAL INFORMATION:  
; APPLICANT: Rodriguez, Moses  
; APPLICANT: Miller, David J.  
; APPLICANT: Asakura, Kunihiko  
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM  
; FILE OF INVENTION: REMYELINATION USING MONOCLONAL AUTOANTIBODIES  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; STREET: Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/779,784  
; FILING DATE: 07-JAN-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/692,084  
; FILING DATE: 08-AUG-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/236,520  
; FILING DATE: 29-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 130 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Mus musculus  
US-08-779-784-35

Query Match 81.2%; Score 601; DB 8; Length 130;  
Best Local Similarity 93.0%; Pred. No. 2e-46;  
Matches 119; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 MGAPAIQLGFLLLFPCTRCDIQMTQSPSSLSASISGQVSVLTCSRASODIGINLHWLOEP 60  
Db 3 MRAPAIQFGFLLLFQGTTCDIQMTQSPSSLSASISGQVSVLTCSRASODIGINLHWLOEP 62  
Qy 61 DGTIKRLIYATSSILGSGVPRKFRSGSDYSLTISSEDFVAYYCLQYASSPYTFGG 120  
Db 63 DGTIKRLIYATSSILGSGVPRKFRSGSDYSLTISSEDFVAYYCLQYASSPYTFGG 122  
Qy 121 GTKLEIKR 128  
Db 123 GTKLEIKR 130

## RESULT 4

US-10-268-883-11  
; Sequence 11, Application US/10268883  
; Publication No. US20030138862A1  
; GENERAL INFORMATION:  
; APPLICANT: Tso, J. Yun  
; APPLICANT: Green, Jennifer Macphate  
; TITLE OF INVENTION: Anti-HLA-DR Antibodies and the Methods of Using Thereof  
; FILE REFERENCE: 05882.0062.NPUS01  
; CURRENT APPLICATION NUMBER: US/10/268,883  
; CURRENT FILING DATE: 2003-03-26  
; PRIOR APPLICATION NUMBER: USSN 60/329,178  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: USSN 60/331,965  
; PRIOR FILING DATE: 2001-11-21  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 130  
; TYPE: PRT  
; ORGANISM: Mouse  
US-10-268-883-11

```

|||||:|:|:|||||:|||||:|||||:|:|:
Db 63 DGTIKRLVYAASLTLDGVPKRFSGSRGSDTSLTISSESEFADYYCLOYTNPLTFGA 122

Qy 121 GTKLEIK 127
    |||||:|
Db 123 GTKLEIK 129

```

KING, DAVID B.  
 TITLE OF INVENTION: Biotransformation of Biotransformation Products For  
 Imaging  
 NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department  
 STREET: Exchange Place, 53 State Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/887,853  
 FILING DATE: 21-Jun-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/133,804  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kelley, Robin D.  
 REGISTRATION NUMBER: 34,637  
 REFERENCE/DOCKET NUMBER: 2054/22  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-248-7477  
 TELEFAX: 617-248-7100  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 243 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
 US-09-887-853-6

OY   81 RESGRSGSDSLTSSLESEDFVAYCLOYASSPYTFGGGTLEIKRAD 130  
DB   194 RESGRSGSDSLTSSLESEDFVVYCLOAIFPPTFGGNTNLEIKRAD 243

RESULT 7  
US-09-924-099-1 ; Sequence 1, Application US/09924099  
; Patent No. US20020128450A1  
; GENERAL INFORMATION:

```

; APPLICANT: NISHIDA, Yoshihiro
; APPLICANT: OKURA, Takanori
; APPLICANT: TANIMOTO, Tadao
; APPLICANT: KURIMOTO, Masashi
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/924,099
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 1
; LENGTH: 108
; TYPE: PPT
; ORGANISM: Mus musculus
US-09-924-099-1
Query Match 68.4%; Score 506; DB 10; Length 108;
Best Local Similarity 92.6%; Pred. No. 4.8e-38;
Matches 100; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Qy 21 DIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQOEPDGTIKRLIYATSSLSGVPK 80
Db 1 DIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQOEPDGTIKRLIYATSSLSGVPK 60
Qy 81 RFGSGRSGSDYSLTISSEDFVYVYCYQYASSPYTFGGGKLEIKR 128
Db 61 RFGSGRSGSDYSLTISSEDFVYVYCYQYASSPYTFGGGKLEIKR 108
RESULT 8
US-09-924-099-9
; Sequence 9, Application US/09924099
; Patent No. US20020128450A1
; GENERAL INFORMATION:
; APPLICANT: NISHIDA, Yoshihiro
; APPLICANT: OKURA, Takanori
; APPLICANT: TANIMOTO, Tadao
; APPLICANT: KURIMOTO, Masashi
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/924,099
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 9
; LENGTH: 237
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially produced peptide in the form of a single
; OTHER INFORMATION: chain
; OTHER INFORMATION: variable region fragment (scFv) which neutralizes IL-18
US-09-924-099-9
Query Match 67.8%; Score 502; DB 10; Length 237;
Best Local Similarity 90.1%; Pred. No. 2.6e-37;
Matches 100; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
Qy 17 GTRCDIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQOEPDGTIKRLIYATSSLSG 76
Db 127 GGGSDIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQOEPDGTIKRLIYATSSLS 186
Qy 77 GVPKRFSGSRSGSDYSLTISSEDFVYVYCYQYASSPYTFGGGKLEIK 127
Db 77 GVPKRFSGSRSGSDYSLTISSEDFVYVYCYQYASSPYTFGGGKLEIK 127

```

```

; APPLICANT: NISHIDA, Yoshihiro
; APPLICANT: OKURA, Takanori
; APPLICANT: TANIMOTO, Tadao
; APPLICANT: KURIMOTO, Masashi
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/924,099
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/338,511
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 289,044/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: JP 365,023/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 10
; LENGTH: 243
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially produced peptide in the form of a single
; OTHER INFORMATION: chain
; OTHER INFORMATION: variable region fragment (scFv) which neutralizes IL-18
US-09-924-099-10
Query Match 67.8%; Score 502; DB 10; Length 243;
Best Local Similarity 90.1%; Pred. No. 2.7e-37;
Matches 100; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
Qy 17 GTRCDIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQOEPDGTIKRLIYATSSLSG 76
Db 127 GGGSDIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQOEPDGTIKRLIYATSSLS 186
Qy 77 GVPKRFSGSRSGSDYSLTISSEDFVYVYCYQYASSPYTFGGGKLEIK 127
Db 187 GVPKRFSGSRSGSDYSLTISSEDFVYVYCYQYASSPYTFGGGKLEIK 237
RESULT 10
US-09-943-906-74
; Sequence 74, Application US/09943906
; Patent No. US20020150571A1
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Williamson, R. Anthony
; APPLICANT: Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/943,906
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>

```



;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 09/550,374  
;; FILING DATE: <Unknown>  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Bozicevic, Karl  
;; REGISTRATION NUMBER: 28,807  
;; REFERENCE/DOCKET NUMBER: 06510/059001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-854-5277  
;; TELEFAX: 415-854-0875  
;; TELEX: <Unknown>  
;; INFORMATION FOR SEQ ID NO: 74:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 109 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 74:  
US-09-943-906-74  
Query Match 67.0%; Score 496; DB 10; Length 109;  
Best Local Similarity 89.0%; Pred. No. 3.8e-37;  
Matches 97; Conservative 7; Mismatches 5; Indels 0; Gaps 0;  
QY 21 DIQMTSPSSLSASLGQVSLTCRASQDIGINLHWLQOEPDGTIKRLIYATSSLSGSGVPK 80  
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
1 ELQMTQPSLSVSLGERSVSLTCRASQDIGSLSNWLQOEPDGTIKRLIYATSSLSGSGVPK 60  
QY 81 RFGSRSGSDYSLTISSESEDFVAYYCLQYASSPYTFGGTKLEIKRA 129  
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
61 RFGSRSGSDYSLTISSESEDLVDYCLQYASSPYTFGGTKLEIKRA 109  
RESULT 11  
US-09-859-053-30  
; Sequence 30, Application US/09859053  
; Patent No. US20020102658A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsuji, Takashi  
; APPLICANT: Tezuka, Katsunari  
; APPLICANT: Hori, No. US20020102658A1uaki  
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST A  
; TITLE OF INVENTION: COSTIMULATORY SIGNAL TRANSDUCTION MOLECULE AILIM AND  
; FILE OF INVENTION: PHARMACEUTICAL USE THEREOF  
; FILE REFERENCE: 06501-079001  
; CURRENT APPLICATION NUMBER: US/09/859,053  
; CURRENT FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: JP 2001-99508  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: JP 2000-147116  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-859-053-30  
Query Match 63.9%; Score 473; DB 10; Length 236;  
Best Local Similarity 66.4%; Pred. No. 1e-34;  
Matches 95; Conservative 15; Mismatches 33; Indels 0; Gaps 0;  
QY 1 MGPAQTLGLLFLPCTRCDIQMTSPSSLSASLGQVSLTCRASQDIGINLHWLQOEP 60  
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
3 MRVPAQLGLLWFFGRCIDIQMTSPSSVSAVGDVITTCRASQGISRLRLAWYQQKP 62  
QY 61 DGTIKRLIYATSSLSGSGSDYSLTISSESEDFVAYYCLQYASSPYTFGG 120  
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
63 GKAPKLLIYVASSLQSGVSRFSGSGGDTFTLTISLPEDATYTCQANSPFTFGQ 122  
QY 121 GTKLEIKRADAPTIVIFPPSSK 143

Db 123 GTKVEIKRTVAAPSVFIFFPSDE 145  
RESULT 12  
US-10-355-780-1  
; Sequence 1, Application US/10355780  
; Publication No. US20030143224A1  
; GENERAL INFORMATION:  
; APPLICANT: Prusiner, Stanley  
; APPLICANT: Safar, Jiri  
; APPLICANT: Williamson, Anthony  
; APPLICANT: Burton, Dennis  
; TITLE OF INVENTION: Antibodies Specific for Ungulate Prp  
; FILE REFERENCE: UCAL-194  
; CURRENT APPLICATION NUMBER: US/10/355,780  
; CURRENT FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: US/09/627,218B  
; PRIOR FILING DATE: 2000-07-27  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 112  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthesized peptide  
US-10-355-780-1  
Query Match 63.6%; Score 471; DB 12; Length 112;  
Best Local Similarity 83.0%; Pred. No. 6.6e-35;  
Matches 93; Conservative 9; Mismatches 10; Indels 0; Gaps 0;  
QY 21 DIQMTSPSSLSASLGQVSLTCRASQDIGINLHWLQOEPDGTIKRLIYATSSLSGSGVPK 80  
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
1 ELVMTQPSLSASLGERSVSLTCRASQDIGNNLWQQKPDGTIKRLIYATSSLSGSGVPK 60  
QY 81 RFGSRSGSDYSLTISSESEDFVAYYCLQYASSPYTFGGTKLEIKRADAA 132  
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
61 RFGSRSGSDYSLTISSESEDFADYVCLQHDFTPLTFGGTKLEIKRTVAA 112  
RESULT 13  
US-09-797-481-2  
; Sequence 2, Application US/09797481  
; Patent No. US20010047083A1  
; GENERAL INFORMATION:  
; APPLICANT: Chatterjee, Malaya  
; Kohler, Heinz  
; Foon, Kenneth A.  
; Chatterjee, Sunil K.  
; TITLE OF INVENTION: MURINE ANTI-IDIOTYPE ANTIBODY 3H1  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/797,481  
; FILING DATE: 28-Feb-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/361,772  
; FILING DATE: 27-JUL-1999  
; APPLICATION NUMBER: US 08/579,940

```
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: US 08/365,484
; FILING DATE: 28-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CATHERINE M. POLIZZI
; REGISTRATION NUMBER: 40,130
; REFERENCE/DOCKET NUMBER: 304142000102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 813-5600
; TELEFAX: (650) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 142 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-797-481-2
;
;
; Query Match
; Best Local Similarity 63.6%; Score 471; DB 9; Length 142;
; Matches 96; Conservative 13; Mismatches 33; Indels 0; Gaps 0;
;
; Qy 1 MGAPAQILGFLLLFPETRCDIOMTQSPSSLSASLGQVSLTCKRASQDIGINLHWLQOEP 60
; Db 1 MVSTAQFLGILLWFPKGIDKMTQSPSSMYASLGERTVITCKASODINGYNWFOQEP 60
;
; Qy 61 DGTIKRLIYATSSILGSGVPRFSGSGSDYSLTISSEDFVAYYCLOYASSPYTFGG 120
; Db 61 GKSPKTLIYRANRLIDGVPFRSGSGGQVYSLTISSEDFVAYYCLOYASSPYTFGG 120
;
; Qy 121 GTKLEIKRAADAAPTIVSIFPPSS 142
; Db 121 GTKLEIKRAADAAPTIVSIFPPSS 142
;
; RESULT 14
; US-09-844-736-4
; Sequence 4, Application US/09844736
; Patent No. US20020041872A1
; GENERAL INFORMATION:
; APPLICANT: Chatterjee, Malaya
; Foon, Kenneth A.
; Chatterjee, Sunil K.
; TITLE OF INVENTION: METHODS OF DELAYING DEVELOPMENT OF
; CEA-ASSOCIATED TUMORS USING ANTI-IDIOTYPE ANTIBODY 3H1
;
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/844,736
; FILING DATE: 09-Apr-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/838,692
; FILING DATE: April 9, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Polizzi, Catherine M.
; REGISTRATION NUMBER: 40,130
; REFERENCE/DOCKET NUMBER: 30414-20004.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
```

```
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 142 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-844-736-4
;
; Query Match
; Best Local Similarity 67.6%; Score 471; DB 9; Length 142;
; Matches 96; Conservative 13; Mismatches 33; Indels 0; Gaps 0;
;
; Qy 1 MGAPAQILGFLLLFPETRCDIOMTQSPSSLSASLGQVSLTCKRASQDIGINLHWLQOEP 60
; Db 1 MVSTAQFLGILLWFPKGIDKMTQSPSSMYASLGERTVITCKASODINGYNWFOQEP 60
;
; Qy 61 DGTIKRLIYATSSILGSGVPRFSGSGSDYSLTISSEDFVAYYCLOYASSPYTFGG 120
; Db 61 GKSPKTLIYRANRLIDGVPFRSGSGGQVYSLTISSEDFVAYYCLOYASSPYTFGG 120
;
; Qy 121 GTKLEIKRAADAAPTIVSIFPPSS 142
; Db 121 GTKLEIKRAADAAPTIVSIFPPSS 142
;
; RESULT 15
; US-10-162-396-4
; Sequence 4, Application US/10162396
; Publication No. US2003007274A1
; GENERAL INFORMATION:
; APPLICANT: Chatterjee, Malaya
; Foon, Kenneth A.
; Chatterjee, Sunil K.
; TITLE OF INVENTION: METHODS OF DELAYING DEVELOPMENT OF
; CEA-ASSOCIATED TUMORS USING ANTI-IDIOTYPE ANTIBODY 3H1
;
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/162,396
; FILING DATE: 31-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/844,736
; FILING DATE: April 27, 2001
; APPLICATION NUMBER: US 08/838,692
; FILING DATE: April 9, 1997
; APPLICATION NUMBER: US 60/044,455
; FILING DATE: April 12, 1996
; APPLICATION NUMBER: US 08/631,085
; FILING DATE: April 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Polizzi, Catherine M.
; REGISTRATION NUMBER: 40,130
; REFERENCE/DOCKET NUMBER: 304142000402
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 4:
```

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 142 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-162-396-4

Query Match          63.6%; Score 471; DB 15; Length 142;
Best Local Similarity 67.6%; Pred. No. 8.6e-35;
Matches 96; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

QY 1 MGAPAIQLGELLFFCTCTCDIOMTQSPSSLSASLGQVSLTCRASODIGINLHWLQOEP 60
Db 1 MVSTAQFLGILLWFFPGIKSDIKMTQSPSSMYASLAGERVITTCASODINGYLNWFQOEP 60

QY 61 DGTIKRLIYATSSILSGGVKRFSGSRGSDYSLTISSESEDFVAYYCLQYASSPYTFGG 120
Db 61 GKSPKTLIYRANRLIDGVPSRFSGSGGQYISLTISSEYEDMGTYCYCLOFDEFPWMFEG 120

QY 121 GTKLEIKRADAAPTVSIFPPSS 142
Db 121 GTKLEIKRADAAPTVSIFPPSS 142
```

Search completed: August 30, 2003, 22:06:56  
Job time : 25.3557 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 22:02:45 ; Search time 17.5168 Seconds  
(without alignments)  
796.063 Million cell updates/sec

Title: us-08-836-455-2  
Perfect score: 740  
Sequence: 1 MGAPAIILGFLLLFPGRNC.....IKRADAPTIVTFPPSKLG 145  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	596	80.5	130	1 KVM5M4	Ig kappa chain pre
2	501	67.7	106	2 PL0260	Ig kappa chain v r
3	497	67.2	106	2 PL0259	Ig kappa chain v r
4	496	67.0	234	2 S14237	Ig kappa chain pre
5	478	64.6	129	2 D32513	Ig kappa chain pre
6	459	62.0	106	2 PL0262	Ig kappa chain v r
7	452	61.1	117	1 KVM53B	Ig kappa chain pre
8	447	60.4	234	2 S01320	Ig kappa chain pre
9	445	60.1	108	4 B47271	nitrophenyl phosph
10	443	59.9	128	2 A26406	Ig kappa chain v r
11	439	59.3	127	2 PH1224	Ig kappa chain pre
12	438	59.2	129	2 S40369	Ig kappa chain - h
13	435	58.8	98	2 PH1062	Ig light chain v r
14	435	58.8	126	2 A34904	Ig kappa chain pre
15	434	58.6	125	2 S40353	Ig kappa chain v-J
16	434	58.6	130	2 S40368	Ig kappa chain - h
17	433	58.5	129	2 B23986	Ig kappa chain pre
18	431	58.2	122	2 A29380	Ig kappa chain v r
19	429	58.0	141	2 A49134	Ig kappa chain pre
20	428.5	57.9	230	2 S33161	Ig kappa chain - s
21	428	57.8	129	2 S40317	Ig kappa chain - h
22	426	57.6	127	2 S52447	Ig kappa chain v r
23	425	57.4	132	2 S40334	Ig kappa chain - h
24	424	57.3	115	2 JL0080	Ig kappa chain pre
25	422	57.0	123	2 S40313	Ig kappa chain v-J
26	419	56.6	88	2 PL0261	Ig kappa chain v r
27	419	56.6	129	2 S52789	Ig kappa chain v r
28	419	56.6	131	2 S40352	Ig kappa chain v-J
29	418	56.5	125	2 S40349	Ig kappa chain v-J

ALIGNMENTS

RESULT 1  
KVM5M4

Ig kappa chain precursor V region (MOPC 41) - mouse  
N:Contains: Ig kappa chain precursor V region VK41  
C:Species: Mus musculus (house mouse)  
C:Date: 24-Apr-1984 #sequence\_revision 24-Apr-1984 #text\_change 28-May-1999  
C:Accession: A93211; B93211; A93815; A94239; A01922; A01923  
R:Seidman, J.G.; Max, E.E.; Leder, P.  
Nature 280, 370-375, 1979

A:Title: A kappa-immunoglobulin gene is formed by site-specific recombination without  
A:Reference number: A93211; MUID:79221900; PMID:111146  
A:Accession: A93211  
A:Molecule type: DNA  
A:Residues: 1-130 <PC41>  
A:Accession: B93211  
A:Molecule type: DNA  
A:Residues: 1-117 <VK41>  
A:Cross-references: GB:V00804; GB:J00566; NID:952127; PIDN:CAA24186.1; PID:g575660  
A>Note: the sequences were determined from the differentiated gene MOPC 41 and the ge  
R:Burstein, Y.; Schechter, I.  
Proc. Natl. Acad. Sci. U.S.A. 74, 716-720, 1977  
A:Title: Amino acid sequence of the NH-2-terminal extra piece segments of the precurs  
A:Reference number: A93815; MUID:77148916; PMID:403522  
A:Accession: A93815  
A:Molecule type: protein  
A:Residues: 1-33 <BUR>  
A>Note: Met-3 is apparently used as an alternative initiator in 25% of the chains  
R:Gray, W.R.; Dreyer, W.J.; Hood, L.  
Science 155, 465-467, 1967  
A:Title: Mechanism of antibody synthesis: size differences between mouse kappa chains  
A:Reference number: A94239; MUID:67056897; PMID:4162931  
A:Accession: A94239  
A:Molecule type: protein  
A:Residues: 23-49, 'B', 51-53, 'LSB', 57-58, '22', 61-62, 'BZ', 65-76, 'B', 78-108, 110-130 <GRA  
A:Experimental source: Bence Jones protein MOPC 41  
C:Genetics:  
A:Introns: 19/1  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (

chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C:Superfamily: Immunoglobulin V region; immunoglobulin homology  
C:Keywords: alternative initiators; heterotetramer; immunoglobulin  
F:1-22/Domain: signal sequence #status experimental <SIG1>  
F:3-22/Domain: signal sequence #status experimental <SIG2>  
F:23-130/Product: Ig kappa chain V region (MOPC 41) #status experimental <M41>  
F:38-112/Domain: immunoglobulin homology <IMM>  
F:45-110/Disulfide bonds: #status predicted

Query Match 80.5%; Score 596; DB 1; Length 130;  
Best Local Similarity 92.2%; Pred. No. 3.7e-42;  
Matches 118; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

0Y 1 MGAPAIILGFLLLFPGRNCDIOMTQSPSSLSASLGORVSLTCRASODIGINLHLWQQEP 60

Db 3 MRAPAIQFGFLLLFQGTCTDIQMTQSPSSLSASLGERVSLTCRASQDIGSSLNWLOQEP 62  
 QY 61 DGTIKRLIYATSSLSGSGVPRFSGSRGSDYSLTISSEDFVAYYCLOYASSPYTFGG 120  
 Db 63 DGTIKRLIYATSSLSGSGVPRFSGSRGSDYSLTISSEDFVAYYCLOYASSPYTFGG 122  
 QY 121 GTKLEIKR 128  
 Db 123 GTKLEIKR 130

## RESULT 2

PL0260  
 Ig kappa chain V region (anti-DNA, DP7VK) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 21-Jan-2000  
 C:Accession: PL0260  
 R:Shlomchik, M.; Mascelli, M.; Shan, H.; Radic, M.Z.; Pisetsky, D.; Marshak-Rothstein, A.  
 J. Exp. Med. 171, 265-297, 1990  
 A:Title: Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic  
 A:Reference number: PL0231; MUID:90111618; PMID:2104919  
 A:Accession: PL0260  
 A:Molecule type: mRNA  
 A:Residues: 1-106 <SHL>  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:1-23/Region: framework 1  
 F:16-90/Domain: immunoglobulin homology <IMM>  
 F:24-34/Region: complementarity-determining 1  
 F:35-49/Region: framework 2  
 F:50-56/Region: complementarity-determining 2  
 F:57-88/Region: framework 3  
 F:89-97/Region: complementarity-determining 3  
 F:98-106/Region: framework 4

Query Match 67.7%; Score 501; DB 2; Length 106;  
 Best Local Similarity 93.4%; Pred. No. 1.9e-34;  
 Matches 99; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 QY 21 DIQMTQSPSSLSASLGERVSLTCRASQDIGINLHWLOQEPDGTIKRLIYATSSLSGSGVPR 80  
 Db 1 DIQMTQSPSSLSASLGERVSLTCRASQDIGSSLNWLOQEPDGTIKRLIYATSSLSGSGVPR 60  
 QY 81 RFGSGRSGSDYSLTISSEDFVAYYCLOYASSPYTFGGGTGLEI 126  
 Db 61 RFGSGRSGSDYSLTISSEDFVAYYCLOYASSPYTFGGGTGLEI 106

## RESULT 3

PL0259  
 Ig kappa chain V region (anti-DNA, DP1LVK) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 21-Jan-2000  
 C:Accession: PL0259  
 R:Shlomchik, M.; Mascelli, M.; Shan, H.; Radic, M.Z.; Pisetsky, D.; Marshak-Rothstein, A.  
 J. Exp. Med. 171, 265-297, 1990  
 A:Title: Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic  
 A:Reference number: PL0231; MUID:90111618; PMID:2104919  
 A:Accession: PL0259  
 A:Molecule type: mRNA  
 A:Residues: 1-106 <SHL>  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:1-23/Region: framework 1  
 F:16-90/Domain: immunoglobulin homology <IMM>  
 F:24-34/Region: complementarity-determining 1  
 F:35-49/Region: framework 2  
 F:50-56/Region: complementarity-determining 2  
 F:57-88/Region: framework 3  
 F:89-97/Region: complementarity-determining 3  
 F:98-106/Region: framework 4

Query Match 67.2%; Score 497; DB 2; Length 106;  
 Best Local Similarity 92.5%; Pred. No. 4.1e-34;  
 Matches 98; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 QY 21 DIQMTQSPSSLSASLGERVSLTCRASQDIGINLHWLOQEPDGTIKRLIYATSSLSGSGVPR 80  
 Db 1 DIQMTQSPSSLSASLGERVSLTCRASQDIGSSLNWLOQEPDGTIKRLIYATSSLSGSGVPR 60  
 QY 81 RFGSGRSGSDYSLTISSEDFVAYYCLOYASSPYTFGGGTGLEI 126  
 Db 61 RFGSGRSGSDYSLTISSEDFVAYYCLOYASSPYTFGGGTGLEI 106

## RESULT 4

SL4237  
 Ig kappa chain precursor (15C5) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 25-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000  
 C:Accession: SL4237  
 R:Vandamme, A.M.; Bullens, F.; Bernar, H.; Nelles, L.; Lijnen, R.H.; Collen, D.  
 Eur. J. Biochem. 192, 767-775, 1990  
 A:Title: Construction and characterization of a recombinant murine monoclonal antibody  
 A:Reference number: SL4236; MUID:91006173; PMID:2209622  
 A:Accession: SL4237  
 A:Molecule type: mRNA  
 A:Residues: 1-234 <VAN>  
 A:Cross-references: EMBL:X56394; NID:g51622; PIDN:CAA39805.1; PID:g51623  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:36-110/Domain: immunoglobulin homology <IMM>

Query Match 67.0%; Score 496; DB 2; Length 234;  
 Best Local Similarity 68.5%; Pred. No. 1.1e-33;  
 Matches 98; Conservative 15; Mismatches 30; Indels 0; Gaps 0;

QY 1 MGAPAIQLGFLLLFPCTRCDIQMTQSPSSLSASLGERVSLTCRASQDIGINLHWLOQEP 60  
 Db 1 MRTPAQLGFLLLFPCTRCDIQMTQSPSSLSASLGERVSLTCRASQDIGINLHWLOQEP 60  
 QY 61 DGTIKRLIYATSSLSGSGVPRFSGSRGSDYSLTISSEDFVAYYCLOYASSPYTFGG 120  
 Db 61 GKSPKTLIYGRNLRVAGVPSRFGSGSGQDYSLTISSEDFVAYYCLOYASSPYTFGG 120  
 QY 121 GTKLEIKRADAAPTTSIFPPSSK 143  
 Db 121 GTKLEIKRADAAPTTSIFPPSE 143

## RESULT 5

D32513  
 Ig kappa chain precursor V region (BXW16) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 21-May-1990 #sequence\_revision 31-Dec-1990 #text\_change 21-Jan-2000  
 C:Accession: D32513  
 R:Kofler, R.; Strohal, R.; Balderas, R.S.; Johnson, M.E.; Noonan, D.J.; Duchosal, M.A.  
 J. Clin. Invest. 82, 852-860, 1988  
 A:Title: Immunoglobulin kappa light chain variable region gene complex organization, a  
 A:Reference number: A94689; MUID:88331394; PMID:3138286  
 A:Accession: D32513  
 A:Molecule type: DNA  
 A:Residues: 1-129 <KOF>  
 A:Cross-references: GB:M20832; NID:g196941; PIDN:AAA38845.1; PID:g196942  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:38-112/Domain: immunoglobulin homology <IMM>

Query Match 64.6%; Score 478; DB 2; Length 129;  
 Best Local Similarity 76.8%; Pred. No. 1.8e-32;  
 Matches 96; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY 3 APAQILGFLLLFPCTRCDIQMTQSPSSLSASLGERVSLTCRASQDIGINLHWLOQEPDG 62  
 Db 5 APAQFLGILLWFPFGARCDIQMTQSPSSMFGSLGDRVSLSCRASQIGIRNLDWYQKPGG 64

KWS53B  
 Ig kappa chain precursor V region (VKM173B) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 29-Jun-1981 #sequence\_revision 29-Jun-1981 #text\_change 21-Jan-2000  
 C:Accession: A01924  
 R:Max, E.B.; Seidman, J.G.; Miller, H.; Leder, P.  
 Cell 21, 793-799, 1980  
 A:Title: Variation in the crossover point of kappa immunoglobulin gene V-J recombination  
 A:Reference number: A01924; MUID:81064681; PMID:6777049  
 A:Accession: A01924  
 A:Molecule type: DNA  
 A:Residues: 1-117 <MAX>  
 A:Cross-references: GB:K00880; NID:g197443; PIDN:AAA39031.1; PID:g197444  
 A:Note: the sequence was determined from the germline gene  
 C:Genetics:  
 A:Introns: 19/1  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into large complexes.  
 C:Superfamily: Immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:1-22/Domain: signal sequence #status predicted <SIG>  
 F:23-117/Product: Ig kappa chain V region (VKM173B) #status predicted <MAT>  
 F:38-112/Domain: immunoglobulin homology <IMM>  
 F:45-110/Disulfide bonds: #status predicted





Search completed: August 30, 2003, 22:08:53  
Job time : 18.5168 secs

## A34904

Query Match 58.8%; Score 435; DB 2; Length 126;  
Best Local Similarity 70.5%; Pred. NO. 5.9e-29;  
Matches 86; Conservative 12; Mismatches 24; Indels 0; Gaps 0;

## RESULT 15

Query Match 58.6%; Score 434; DB 2; Length 125;  
Best Local Similarity 68.0%; Pred. No. 7.1e-29;  
Matches 85; Conservative 14; Mismatches 26; Indels 0; Gaps 0;

QY	7	ILGFLLLFPPTCRDIQMTQSPSSLSASLGQRVSLTCSRQSDIGINLHWLQQPDGPKIR	66
Db	1	LGFLLLWFPGARCDIQMTQSPSSLSASVGRVITTCRASQIGNDLQWYQKPGKAPR	60
QY	67	LIYATSSLGSGVPRFSGSRGSDYSITLSLESEDFVAYYICQYASSPYTFGGGKLEI	126
Db	61	LIYAASSFQSGVPRFSGSGTGFTLTISGLQPEDFATYYICLQHSNYPITFGGGTRVEI	120
QY	127	KRAA	131
Db	121	KRTVA	125

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 19:57:10 ; Search time 10.2181 Seconds  
(without alignments)  
667.331 Million cell updates/sec

Title: US-08-836-455-2

Perfect score: 740

Sequence: 1 MGAPQILGFLLLFPGRRC.....IKRADAAPTVSIFPPSKLG 145

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	596	80.5	130	1	KV5G_MOUSE
2	452	61.1	117	1	KV5H_MOUSE
3	412	55.7	129	1	KV5E_MOUSE
4	394	53.2	129	1	KV1W_HUMAN
5	389	52.6	108	1	KV5J_MOUSE
6	384	51.9	108	1	KV5K_MOUSE
7	383	51.8	108	1	KV5Q_MOUSE
8	382	51.6	108	1	KV5M_MOUSE
9	382	51.6	108	1	KV5N_MOUSE
10	381	51.5	117	1	KV1L_HUMAN
11	379	51.2	108	1	KV5L_MOUSE
12	373	50.4	108	1	KV1R_HUMAN
13	370	50.0	115	1	KV5F_MOUSE
14	368	49.7	129	1	KV1X_HUMAN
15	367	49.6	108	1	KV5S_MOUSE
16	365	49.3	108	1	KV5U_MOUSE
17	364	49.2	108	1	KV5Q_MOUSE
18	363	49.1	108	1	KV5T_MOUSE
19	362	48.9	117	1	KV1J_HUMAN
20	354	47.8	108	1	KV1O_HUMAN
21	353.5	47.8	129	1	KV3H_HUMAN
22	353.5	47.7	129	1	KV3M_HUMAN
23	353	47.7	108	1	KV5R_MOUSE
24	352.5	47.6	129	1	KV3L_HUMAN
25	351	47.4	108	1	KV1Y_HUMAN
26	350	47.3	128	1	KV3K_HUMAN
27	348	47.0	108	1	KV1G_HUMAN
28	348	47.0	108	1	KV1L_HUMAN
29	348	47.0	108	1	KV1P_HUMAN
30	346	46.8	108	1	KV1H_HUMAN
31	346	46.8	108	1	KV1V_HUMAN
32	341	46.1	108	1	KV1B_HUMAN
33	339	45.8	108	1	KV1C_HUMAN

## ALIGNMENTS

### RESULT 1

ID	KV5G_MOUSE	STANDARD;	PRT;	130 AA.
AC	P01639; P01640;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DE	Ig kappa chain V-V region MOPC 41 precursor.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=79221900; PubMed=111146;			
RA	Seidman J.G., Max E.E., Leder P.;			
RT	"A kappa-immunoglobulin gene is formed by site-specific recombination			
RT	without further somatic mutation.";			
RL	Nature 280:370-375(1979).			
RN	[2]			
RP	SEQUENCE OF 1-33.			
RX	MEDLINE=77148916; PubMed=403522;			
RA	Burstein Y., Schechter I.;			
RT	"Amino acid sequence of the NH2-terminal extra piece segments of the			
RT	precursors of mouse immunoglobulin lambdaI-type and kappa-type light			
RL	chains.";			
RN	[3]			
RP	SEQUENCE OF 23-130.			
RX	MEDLINE=67056897; PubMed=4162931;			
RA	Gray W.R., Dreyer W.J., Hood L.;			
RT	"Mechanism of antibody synthesis: size differences between mouse			
RT	kappa chains.";			
RL	Science 155:465-467(1967).			
CC	-I- MISCELLANEOUS: THIS PRECURSOR WAS SYNTHESIZED IN A CELL-FREE			
CC	SYSTEM DIRECTED BY MENA ISOLATED FROM MYELOMA POLYSOMES.			
CC	-I- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.			
DR	PIR; A93211; KVM5M4.			
DR	HSSP; P01607; IREI.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	InterPro; IPR003596; Ig_V.			
DR	Pfam; PF00047; Ig; 1.			
DR	SMART; SM00406; IGV; 1.			
DR	PROSITE; PS50835; IG_LIKE; 1.			
KW	Immunoglobulin V region; Signal; Bence-Jones protein.			
FT	SIGNAL 1 22			
FT	CHAIN 23 130			
FT	DOMAIN 23 45			
FT	DOMAIN 46 56			
FT	DOMAIN 57 71			
FT	DOMAIN 72 78			
FT	DOMAIN 79 110			
FT	DOMAIN 111 119			
FT	DOMAIN 120 129			
FT	DISULFID 45 110			

P01597 homo sapien  
P01606 homo sapien  
P01596 homo sapien  
P01611 homo sapien  
P01593 homo sapien  
P01649 mus musculu  
P01612 homo sapien  
P01605 homo sapien  
P06310 homo sapien  
P01603 homo sapien  
P01598 homo sapien  
P01609 homo sapien

KV1E\_HUMAN 108 1  
KV1D\_HUMAN 108 1  
KV1S\_HUMAN 107 1  
KV1A\_HUMAN 108 1  
KV5P\_MOUSE 108 1  
KV1T\_HUMAN 109 1  
KV1M\_HUMAN 108 1  
KV2F\_HUMAN 133 1  
KV1K\_HUMAN 108 1  
KV1E\_HUMAN 108 1  
KV1Q\_HUMAN 108 1

34 339 45.8  
35 339 45.8  
36 338.5 45.7  
37 338 45.7  
38 337 45.5  
39 332 44.9  
40 331.5 44.8  
41 331 44.7  
42 330.5 44.7  
43 327 44.2  
44 326 44.1  
45 325 43.9

FT VARIANT 1 2 MISSING (IN 25% OF THE MOLECULES).

FT NON\_TER 130 130

SQ SEQUENCE 130 AA; 14311 MW; 5EPE0FE71D5F1BEC CRC64;

Query Match 80.5%; Score 596; DB 1; Length 130;

Best Local Similarity 92.2%; Pred. No. 2.9e-52;

Matches 118; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 MGAPAIQILGFLLLFPPTGTRCDIQMTQSPSSLSASLGQVSLTCRASQDIGINLHWLQOEP 60

DB 3 MRAPAIQIFGFLLLFPPTGTRCDIQMTQSPSSLSASLGQVSLTCRASQDIGINLHWLQOEP 62

QY 61 DGTIKRLIYATSSLSGSGVPRKFRSGSDYSLTISSESEDFVAYYCLQYASSPYTFGG 120

DB 63 DGTIKRLIYATSSLSGSGVPRKFRSGSDYSLTISSESEDFVAYYCLQYASSPYTFGG 122

QY 121 GTKLEIKR 128

DB 123 GTKLEIKR 130

#### RESULT 2

KV5H\_MOUSE

ID KV5H\_MOUSE STANDARD; PRT; 117 AA.

AC P01641;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig kappa chain V-V region MOPC 173B precursor.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=81064681; PubMed=6777049;

RA Max E.E., Seidman J.G., Miller H., Leder P.;

RT "Variation in the crossover point of kappa immunoglobulin gene V-J

recombination: evidence from a cryptic gene.";

RL Cell 21:793-799(1980).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; K00880; AAA39031.1; -.

DR PIR; A01924; KVM53B.

DR HSSP; P01607; IREI.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR003596; Ig\_V.

DR Pfam; PF00047; Ig\_1.

DR SMART; SM00406; Igv; 1.

KW Immunoglobulin V region; Signal.

FT SIGNAL 1 22

FT CHAIN 23 117 IG KAPPA CHAIN V-V REGION MOPC 173B.

FT DOMAIN 23 45 FRAMEWORK-1.

FT DOMAIN 23 45 COMPLEMENTARITY-DETERMINING-1.

FT DOMAIN 46 56 FRAMEWORK-2.

FT DOMAIN 57 71 FRAMEWORK-3.

FT DOMAIN 72 78 COMPLEMENTARITY-DETERMINING-2.

FT DOMAIN 79 110 FRAMEWORK-3.

FT DOMAIN 111 >117 COMPLEMENTARITY-DETERMINING-3.

FT DISULFID 45 110 BY SIMILARITY.

FT NON\_TER 117 117

SQ SEQUENCE 117 AA; 12954 MW; 24B3D4B9AC2EAD6C CRC64;

Query Match

Best Local Similarity 61.1%; Score 452; DB 1; Length 117;

Matches 92; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

QY 1 MGAPAIQILGFLLLFPPTGTRCDIQMTQSPSSLSASLGQVSLTCRASQDIGINLHWLQOEP 60

DB 3 MRAPAVGFGFLLLFPPTGTRCDIQMTQSPSSLSASLGQVSLTCRASQDIGINLHWLQOEP 62

QY 61 DGTIKRLIYATSSLSGSGVPRKFRSGSDYSLTISSESEDFVAYYCLQYASSP 115

DB 63 GETIKHLIYETNSLDGSGVPRKFRSGSDYSLTISSESEDFVAYYCLQYASSP 117

#### RESULT 3

KV5E\_MOUSE

ID KV5E\_MOUSE STANDARD; PRT; 128 AA.

AC P01637;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig kappa chain V-V region T1 precursor.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=81052342; PubMed=6776411;

RA Altenburger W., Steinmetz M., Zachau H.G.;

RT "Functional and non-functional joining in immunoglobulin light chain

genes of a mouse myeloma.";

RL Nature 287:603-607(1980).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; V00772; CAA24150.1; -.

DR PIR; A01920; KVM5T1.

DR HSSP; P80362; IWT1.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR003596; Ig\_V.

DR Pfam; PF00047; Ig\_1.

DR SMART; SM00406; Igv; 1.

DR PROSITE; PS50835; IG\_LIKE; 1.

KW Immunoglobulin V region; Signal.

FT SIGNAL 1 20

FT CHAIN 21 128 IG KAPPA CHAIN V-V REGION T1.

FT DOMAIN 21 43 FRAMEWORK-1.

FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.

FT DOMAIN 55 69 FRAMEWORK-2.

FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.

FT DOMAIN 77 108 FRAMEWORK-3.

FT DOMAIN 109 117 COMPLEMENTARITY-DETERMINING-3.

FT DOMAIN 118 127 FRAMEWORK-4.

FT DISULFID 43 108 BY SIMILARITY.

FT NON\_TER 128 128

SQ SEQUENCE 128 AA; 14385 MW; AFA563D31BB7E05 CRC64;

Query Match

Best Local Similarity 55.7%; Score 412; DB 1; Length 128;

Matches 81; Conservative 14; Mismatches 33; Indels 0; Gaps 0;

QY 1 MGAPAIQILGFLLLFPPTGTRCDIQMTQSPSSLSASLGQVSLTCRASQDIGINLHWLQOEP 60

DB 1 MRTPAQFLGILLFPPTGTRCDIKMTQSPSSNYSASLGQVSLTCRASQDINSYLTWFQKP 60

QY 61 DGTIKRLIYATSSLSGSGVPRKFRSGSDYSLTISSESEDFVAYYCLQYASSPYTFGG 120

DB 61 GKSPKTLIYRANRLVDGVPFRSGSGQDFSLTISSESEDFVAYYCLQYDEFLPTFGA 120

QY 121 GTKLEIKR 128

Db 121 GTKLELR 128

## RESULT 4

ID KVIW\_HUMAN STANDARD; PRT; 129 AA.  
 AC P04431;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig kappa chain V-I region Walker precursor.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RZ SEQUENCE FROM N.A.  
 RA MEDLINE=85014148; PubMed=6091049;  
 RX Klobbeck H.G., Combratio G., Zachau H.G.;  
 RT "Immunoglobulin genes of the kappa light chain type from two human  
 lymphoid cell lines are closely related.";  
 RL Nucleic Acids Res. 12:6995-7006(1984).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X00965; CAA25477.1; ALT\_TERM.  
 DR PIR; A01883; KIHUWK.  
 DR HSP; P01607; IREI.  
 DR GO; GO:0003576; C:extracellular; NAS.  
 DR GO; GO:0003823; F:antigen binding activity; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR InterPro; IPR007110; Ig-Like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig\_1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Immunoglobulin V region; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 129 IG KAPPA CHAIN V-I REGION WALKER.  
 FT DOMAIN 23 45 FRAMEWORK-1.  
 FT DOMAIN 46 56 COMPLEMENTARITY-DETERMINING-1.  
 FT DOMAIN 57 71 FRAMEWORK-2.  
 FT DOMAIN 72 78 COMPLEMENTARITY-DETERMINING-2.  
 FT DOMAIN 79 110 FRAMEWORK-3.  
 FT DOMAIN 111 119 COMPLEMENTARITY-DETERMINING-3.  
 FT DOMAIN 120 129 FRAMEWORK-4.  
 FT DISULFID 45 110 BY SIMILARITY.  
 FT NON\_TER 129 129  
 SQ SEQUENCE 129 AA; 144069 MW; F941FA07D4AFC2F9 CRC64;

## RESULT 5

```

KV5J_MOUSE
ID AC KV5J_MOUSE STANDARD: PRT; 108 AA.
IC P01643;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region MOPC 173.
DE Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=76091934; PubMed=812696;
RA Schiff C., Fougereau M.;
RT *Determination of the primary structure of a mouse IgG2a
RT immunoglobulin. Amino-acid sequence of the light chain.*;
RL Eur. J. Biochem. 59:525-537(1975).
CC -I- MISCLEANEORIS. THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN
DR PIR; A01926; KVS73.
DR HSP; P01607; IREI.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
KW Immunoglobulin V region.
KW DOMAIN 1 23
FT DOMAIN 24 34 FRAMEWORK-1.
FT DOMAIN 35 49 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 50 56 FRAMEWORK-2.
FT DOMAIN 57 88 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 89 97 FRAMEWORK-3.
FT DOMAIN 98 108 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 23 88 FRAMEWORK-4.
FT NON_TER 108 108 BY SIMILARITY.
SQ SEQUENCE 108 AA; 11819 MW; 2AD29D92A72AA03 CRC64;

Query Match 52.6%; Score 389; DB 1; Length 108;
Best Local Similarity 69.4%; Pred. No. 8e-32;
Matches 75; Conservative 11; Mismatches 22; Indels 0; Gaps 0

QY 21 DIQMTQSPSSLSASLGQRVSLTCSRASQIGINHLWLQEPDGTIKRLIYATSSLGSC
Db 1 DIQMTQTSSLSASLGDRVTISCSASQISGNVLBWYQKPGDTVKLLIYTTSLHSC
QY 81 RFGSGRSQSYSLTISLESSEDVYAYCQYASSYTFGGGKLEIKR 128
Db 61 RFGSGSGSDYSLTISLSPZPIATYICQYQSKLPTFGGKLEIKR 108

```

## RESULT 6

KV5K_MOUSE	KV5K_MOUSE	STANDARD;	PRT;	108 AA.
ID	KV5K_MOUSE			
AC	P01644;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Ig kappa chain V-V region HP R15.7.			
DE	Mus musculus (Mouse).			
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE.			
RC	STRAIN=A/J;			
RC	MEDLINE=82150934; PubMed=6801658;			
RT	Siegelman M., Capra J.D.;			
RA	"Complete amino acid sequence of light chain variable regions deri			

Program region.	
1	23
DOMAIN	FRAMEWORK-1.
FT	COMPLEMENTARITY-DETERMINING-1.
24	FRAMEWORK-2.
DOMAIN	COMPLEMENTARITY-DETERMINING-2.
34	COMPLEMENTARITY-DETERMINING-3.
FT	FRAMEWORK-3.
50	COMPLEMENTARITY-DETERMINING-3.
DOMAIN	
57	
FT	
89	
DOMAIN	
97	

Program region.	
1	23
DOMAIN	FRAMEWORK-1.
FT	COMPLEMENTARITY-DETERMINING-1.
24	FRAMEWORK-2.
DOMAIN	COMPLEMENTARITY-DETERMINING-2.
34	COMPLEMENTARITY-DETERMINING-3.
FT	FRAMEWORK-3.
50	COMPLEMENTARITY-DETERMINING-3.
DOMAIN	
57	
FT	
89	
DOMAIN	
97	

```

81 RFSGRSGSDYSLTISSESEDFVAYYCLQYASSPYTFGGGKLEIKR 128
      ||||| : ||||| : ||| : ||||| : ||||| : ||||| : ||||| : ||||| :
61 RFSGSGGSDYSLTISNLEQEDATVFCQQGYMLPRTFGGKLEIKR 108

```

```

KV5N_MOUSE
ID KV5N_MOUSE STANDARD; PRT; 108 AA.
AC P01647;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ig kappa chain V-V region HP 124E1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC STRAIN=A/J;
RX MEDLINE=82150934; PubMed=6801658;
RA Siegelman M., Capra J.D.;
RT "Complete amino acid sequence of light chain variable regions derived
RT from five monoclonal anti-p-azophenylarsonate antibodies differing
RT with respect to a crossreactive idiotype.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:7679-7683(1981).
CC -1- MISCELLANEOUS: ANTI-ARSONATE HYBRIDOMA PROTEIN.
DR HSSP; P01607; 1REI.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 34
FT DOMAIN 35 49
FT DOMAIN 50 56
FT DOMAIN 57 88
FT DOMAIN 89 97
FT DOMAIN 98 108
FT DISULFID 23 88
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11965 MW; 39971BC653EFEFA2 CRC64;

Query Match 51.6%; Score 382; DB 1; Length 108;
Best Local Similarity 70.4%; Pred. No. 4e-31;
Matches 76; Conservative 11; Mismatches 21; Indels 0; Gaps 0;

QY 21 DIQMTQSPSSLSASLGQRVSLTCSRASQDIGINLHWLQEPDGTIKRLIYATSLGSGVPK 80
DB 1 DIQMTQTTSSLSASLGDRVTISCRASQDINNLYNQKPDGTGTVKLLIYTSKRLHSGVPS 60

QY 81 RESGSRSGSDYSLTISSESEDFVAYYCYQYASSPYTFGGGKLEIKR 128
DB 61 RFSGSGGPDYSLTISLNLEQEDVATYFCQGGKTLPRTFGGGKLEIKR 108

RESULT 10
KV5L_MOUSE
ID KV5L_MOUSE STANDARD; PRT; 117 AA.
AC P01601;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region HK101 precursor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81098966; PubMed=6779204;
RA Bentley D.L., Rabbitts T.H.;
RT "Human immunoglobulin variable region genes -- DNA sequences of two V
RT kappa genes and a pseudogene.";
RL Nature 288:730-733(1990).
RN [2]

SEQUENCE FROM N.A.
RP MEDLINE=83129397; PubMed=6402305;
RA Bentley D.L., Rabbitts T.H.;
RT "Evolution of immunoglobulin V genes: evidence indicating that
RT recently duplicated human V kappa sequences have diverged by gene
RT conversion.";
RL Cell 32:181-189(1983).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K01322; AAA58930.1; -
CC EMBL; K01324; AAA58932.1; -
CC EMBL; V00558; CAA23824.1; -
CC PIR; A01881; K1H011.
CC HSSP; P01607; 1REI.
CC GO; GO:0005576; C:extracellular; NAS.
CC GO; GO:0003823; F:antigen binding activity; NAS.
CC GO; GO:0006955; P:immune response; NAS.
CC InterPro; IPR007110; Ig-Like.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003596; Ig_V.
CC Pfam; PF00047; Ig; 1.
CC SMART; SM00406; IGV; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 22
FT CHAIN 23 >117 IG KAPPA CHAIN V-I REGION HK101.
FT DOMAIN 23 45 FRAMEWORK-1.
FT DOMAIN 46 56 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 57 71 FRAMEWORK-2.
FT DOMAIN 72 78 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 79 110 FRAMEWORK-3.
FT DOMAIN 111 >117 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 45 110 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12799 MW; D7D0FF3718CEF587 CRC64;

Query Match 51.5%; Score 381; DB 1; Length 117;
Best Local Similarity 68.5%; Pred. No. 5.5e-31;
Matches 76; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

QY 5 AQILGFLLLFPCTRCDIQMTQSPSSLSASLGQRVSLTCSRASQDIGINLHWLQEPDGTI 64
DB 7 AQLGLLLLCFFPGARCDIQMTQSPSSLSASVGRVITTCRARGISGLWLAAYQOKPEKAP 66

QY 65 KRLIYATSLGSGVPRKFRSGSRGSDYSLTISSESEDFVAYYCYQYASSP 115
DB 67 KSLIYAASSLQSGVPRFSGSGSGTDTLTLSLQPEDFATYYCQYNSYP 117

RESULT 11
KV5L_MOUSE
ID KV5L_MOUSE STANDARD; PRT; 108 AA.
AC P01645;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ig kappa chain V-V region HP 93G7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC STRAIN=A/J;
RX MEDLINE=82150934; PubMed=6801658;
RA Siegelman M., Capra J.D.;

```

\*Complete amino acid sequence of light chain variable regions derived from five monoclonal anti-p-azophenylarsenate antibodies differing with respect to a crossreactive idiotype.\*;  
 Proc. Natl. Acad. Sci. U.S.A. 78:7679-7683(1981).  
 CC -1- MISCELLANEOUS: ANTI-ARSONATE HYBRIDOMA PROTEIN.  
 DR HSP; P01607; IREI.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 DR PROSITE: PS00835; IG\_LIKE; 1.  
 DR Immunoglobulin V region.

FT DOMAIN 1 23  
 FT DOMAIN 24 34  
 FT DOMAIN 35 49  
 FT DOMAIN 50 56  
 FT DOMAIN 57 88  
 FT DOMAIN 57 88  
 FT DOMAIN 89 97  
 FT DOMAIN 98 108  
 FT DISULFID 23 88  
 FT NON\_TER 108 108  
 SQ SEQUENCE 108 AA; 11954 MW; 22F4642C53EFF58E CRC64;  
 Query Match 51.2%; Score 379; DB 1; Length 108;  
 Best Local Similarity 70.4%; Pred. No. 7.9e-31;  
 Matches 76; Conservative 10; Mismatches 22; Indels 0; Gaps 0;

QY 21 DIQMTSPSSLSASLGORVSLTCTCRASODIGINLHWLQOEPDGTIKRLIYATSSIGSGVPK 80  
 DB 1 DIQMTSSLSASLGORVSLTCTCRASODISNYLNWYQKPDGIVKLIIYTSRLHSGVPS 60  
 QY 81 RFGSGSGSDYSLTISSESEDFVAYYCLQYASSPYTFGGGKLEIKR 128  
 DB 61 RFGSGSGSDYSLTISNLEQDIATYPCQGNMLPRTFGGGKLEIKR 108

## RESULT 12

KVIR\_HUMAN  
 ID KVIR\_HUMAN STANDARD; PRT; 108 AA.  
 AC P01610;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig kappa chain V-I region WEA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=83273707; PubMed=6410398;  
 RA Goni F., Frangione B.;  
 RT "Amino acid sequence of the Fv region of a human monoclonal IgM (protein WEA) with antibody activity against 3,4-pyruvylated galactose in Klebsiella polysaccharides K30 and K33.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4837-4841(1983).  
 CC -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM A MONOCLONAL ANTIBODY AGAINST 3,4-PYRUVYLATED GALACTOSE AND ISOLATED FROM A PATIENT WITH WALDENSTROM'S MACROGLOBULINEMIA.  
 CC PIR: A01876; KIHUWE.  
 DR HSP; P80362; IWL.  
 DR GO: GO:0005576; C:extracellular; NAS.  
 DR GO: GO:0003823; F:antigen binding activity; NAS.  
 DR GO: GO:0006955; P:immune response; NAS.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 DR PROSITE: PS00835; IG\_LIKE; 1.  
 DR Immunoglobulin V region; Monoclonal antibody.  
 KW IMUNOGLOBULIN V region; Monoclonal antibody.  
 FT DOMAIN 1 23

FT DOMAIN 24 34  
 FT DOMAIN 35 49  
 FT DOMAIN 50 56  
 FT DOMAIN 57 88  
 FT DOMAIN 89 97  
 FT DOMAIN 98 107  
 FT DISULFID 23 88  
 FT NON\_TER 108 108  
 SQ SEQUENCE 108 AA; 11840 MW; 9249B61F0945618C CRC64;  
 Query Match 50.4%; Score 373; DB 1; Length 108;  
 Best Local Similarity 66.7%; Pred. No. 3.1e-30;  
 Matches 72; Conservative 16; Mismatches 20; Indels 0; Gaps 0;

QY 21 DIQMTSPSSLSASLGORVSLTCTCRASODIGINLHWLQOEPDGTIKRLIYATSSIGSGVPK 80  
 DB 1 DIQMTSPSSLSASVGDVITCTCRASQIRNDLTWYQKPGCTAPKRLIYCATSLQSGVPS 60  
 QY 81 RFGSGSGSDYSLTISSESEDFVAYYCLQYASSPYTFGGGKLEIKR 128  
 DB 61 RFGSGSGTEFTLTINSLOPDEATYCYLQYSSFPPTFGQGTKEVVKR 108

## RESULT 13

KVSF\_MOUSE  
 ID KVSF\_MOUSE STANDARD; PRT; 115 AA.  
 AC P01638;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig kappa chain V-V region L6 precursor (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=81220975; PubMed=6264318;  
 RA Pech M., Hochtl J., Schnell H., Zachau H.G.;  
 RT "Differences between germ-line and rearranged immunoglobulin V kappa coding sequences suggest a localized mutation mechanism.";  
 RL Nature 291:668-670(1981).  
 DR PIR: A01921; KMSL6.  
 DR HSP; P01607; IREI.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 DR PROSITE: PS00835; IG\_LIKE; 1.  
 KW Immunoglobulin V region; Signal.  
 FT SIGNAL 1 20  
 FT CHAIN 21 >115 IG KAPPA CHAIN V-V REGION L6.  
 FT DOMAIN 21 43  
 FT DOMAIN 44 54  
 FT DOMAIN 55 69  
 FT DOMAIN 70 76  
 FT DOMAIN 77 108  
 FT DOMAIN 109 >115  
 FT DISULFID 43 108  
 FT NON\_TER 115 115  
 SQ SEQUENCE 115 AA; 12986 MW; BA852C58F328E1C3 CRC64;  
 Query Match 50.0%; Score 370; DB 1; Length 115;  
 Best Local Similarity 64.3%; Pred. No. 6.7e-30;  
 Matches 74; Conservative 10; Mismatches 31; Indels 0; Gaps 0;

QY 1 MGAPAQILGLLLLPCTRCDIQMTSPSSLSASLGORVSLTCTCRASODIGINLHWLQOEP 60  
 DB 1 MRPAPQFLGLLLLPCTRCDIQMTSPSSLSASLGORVSLTCTCRASODIGINLHWLQOEP 60  
 QY 61 DGTIKRLIYATSSIGSGVPSRFGSGSGSDYSLTISSESEDFVAYYCLQYASSP 115



Db 123 GTKVDNK 129

Job time : 10.2181 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 21:05:00 ; Search time 34.547 Seconds  
(without alignments)  
1083.094 Million cell updates/sec

Title: US-08-836-455-2  
Perfect score: 740  
Sequence: 1 MGAQAIIIGLFLLLFPGRFC.....IKRADAPTSTIFPSSKLG 149

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	642	86.8	127	11 Q925S9	Q925S9 mus musculus
2	513	69.3	234	11 Q8R062	Q8R062 mus musculus
3	506	68.4	234	11 Q91WF8	Q91WF8 mus musculus
4	492	66.5	233	11 Q91WS9	Q91WS9 mus musculus
5	450	60.8	234	11 Q8VCP0	Q8VCP0 mus musculus
6	433	58.5	214	11 Q9RIA5	Q9RIA5 mus musculus
7	416	56.2	234	4 Q8NEK1	Q8NEK1 homo sapien
8	391.5	52.9	239	11 Q8K0F8	Q8K0F8 mus musculus
9	382.5	51.7	235	11 Q91W12	Q91W12 mus musculus
10	380	51.4	107	11 Q9UL84	Q9UL84 mus musculus
11	374	50.5	238	11 Q99M37	Q99M37 mus musculus
12	371.5	50.2	239	4 Q8TCD0	Q8TCD0 homo sapien
13	370	50.0	238	11 Q8VCI6	Q8VCI6 mus musculus
14	363.5	49.1	134	11 Q8VDD0	Q8VDD0 mus musculus
15	360.5	48.7	114	11 Q8K1F1	Q8K1F1 mus musculus
16	360	48.6	108	4 Q9UL70	Q9UL70 homo sapien

17	359.5	48.6	107	4 Q96SA9	Q96SA9 homo sapien
18	357.5	48.3	112	11 Q8K1F3	Q8K1F3 mus musculus
19	357	48.2	108	4 Q9UL77	Q9UL77 homo sapien
20	356.5	48.2	239	11 Q8VC55	Q8VC55 mus musculus
21	352.5	47.6	239	4 Q8NEK0	Q8NEK0 homo sapien
22	346	46.8	234	11 Q8R028	Q8R028 mus musculus
23	342.5	46.3	112	11 Q8K1F2	Q8K1F2 mus musculus
24	341.5	46.1	112	11 Q8K1F0	Q8K1F0 mus musculus
25	339	45.8	108	11 Q8VIJ0	Q8VIJ0 mus musculus
26	338	45.7	116	4 Q96PF6	Q96PF6 homo sapien
27	337.5	45.6	107	4 Q9UL81	Q9UL81 homo sapien
28	336	45.4	108	4 Q9UL79	Q9UL79 homo sapien
29	329	44.5	298	11 Q9QYF0	Q9QYF0 mus musculus
30	328	44.3	109	11 Q920E6	Q920E6 mus musculus
31	317.5	42.9	109	4 Q9UL78	Q9UL78 homo sapien
32	316	42.7	108	4 Q9UL83	Q9UL83 homo sapien
33	315.5	42.6	109	4 Q9UL85	Q9UL85 homo sapien
34	310	41.9	111	11 Q920E9	Q920E9 mus musculus
35	299.5	40.5	109	4 Q9UL86	Q9UL86 homo sapien
36	296.5	40.1	106	5 Q9UL80	Q9UL80 schistosoma
37	295	39.9	97	11 Q9JL76	Q9JL76 mus musculus
38	294.5	39.8	241	11 Q921A6	Q921A6 mus musculus
39	293	39.6	101	11 Q9JL78	Q9JL78 mus musculus
40	275	37.2	114	4 Q9UL80	Q9UL80 homo sapien
41	273	36.9	236	4 Q8NEJ1	Q8NEJ1 homo sapien
42	270.5	36.6	237	4 Q8WUK4	Q8WUK4 homo sapien
43	269	36.4	103	11 Q9JL80	Q9JL80 mus musculus
44	268.5	36.3	237	4 Q8WTU6	Q8WTU6 homo sapien
45	267	36.1	107	11 Q9ER29	Q9ER29 mus musculus

ALIGNMENTS

RESULT 1

Q925S9 ID Q925S9 PRELIMINARY; PRT; 127 AA.  
AC Q925S9;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Immunoglobulin light chain (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c;  
RX MEDLINE=99306687; PubMed=10380019;  
RA Tripathi P.K., Qin H., Bhattacharya-Chatterjee M., Ceriani R.L.,  
RA Foon K.A., Chatterjee S.K.;  
RT "Construction and characterization of a chimeric fusion protein  
RT consisting of an anti-idiotypic antibody mimicking a breast cancer-  
RT associated antigen and the cytokine GM-CSF";  
RL Hybridoma 18:193-202(1999).  
DR EMBL; AF124721; AAK55120.1;  
DR InterPro: IPR007110; Ig-like.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IG\_LIKE; 1.  
FT NON\_TER 127  
SQ SEQUENCE 127 AA; 13794 MW; 13F61BEBB981FA5 CRC64;

Query Match 86.8%; Score 642; DB 11; Length 127;  
Best Local Similarity 99.2%; Pred. No. 8.5e-60;  
Matches 126; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGAQAIIIGLFLLLFPGRFCDIQMTQSPSSLSASLGQRVSLTCRASODIGINLHWLQOEP 60  
DB 1 MRPAQIIIGLFLLLFPGRFCDIQMTQSPSSLSASLGQRVSLTCRASODIGINLHWLQOEP 60





```

DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031498; AAH31498.1; -
DR InterPro; IPR002198; ADH_Short.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; Igc1; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00061; ADH_SHORT; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 239 AA; 26366 MW; D7BE84398AA341F0 CRC64;

Query Match 52.9%; Score 391.5; DB 11; Length 239;
Best Local Similarity 55.4%; Pred. No. 4.7e-33;
Matches 82; Conservative 20; Mismatches 41; Indels 5; Gaps 1;

QY 1 MGAPAILGLFLLFPFG--TRCDIQMTQSPSSLSASLGQVSLTCTCRASQDIGNLHWLQ 55
DB 1 MMSPAFLFLLVLSQINGDVMTQTPTFLSVITIGPASIKSSQSLSFYTNKWLW 60

QY 56 LQEPDGTIKRLIYATSSLSGSGVSKRFSRSGDYSLTISSEDFVAYYCLOYASSP 115
DB 61 LLQRCQSPKRLISLVSKLDGVPDRFSGSGGTDTFTLKISRVEAEDLGVYICLQSTHP 120

QY 116 YTFGGTGKLEIKRADAAPTVSIFPPSSK 143
DB 121 YTFGGTGKLEIKRADAAPTVSIFPPSS 148

RESULT 9
Q91W12 PRELIMINARY; PRT; 235 AA.
ID Q91W12
AC Q91W12
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Unknown (Protein for MGC:6582).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006643; AAH06643.1; -
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 235 AA; 26021 MW; 5FC73BDEBD5E8FEF CRC64;

Query Match 51.7%; Score 382.5; DB 11; Length 235;
Best Local Similarity 55.9%; Pred. No. 4e-32;
Matches 81; Conservative 22; Mismatches 39; Indels 3; Gaps 2;

```

```

QY 1 MGAPAILGLFLLFPFG--TRCDIQMTQSPSSLSASLGQVSLTCTCRASQDIGNLHWLQ 58
DB 1 MDFQVQIFSLTISASVILSRGQIVLTQSPAINSPGERVTMTCSASSVS-HMHWQQ 59

QY 59 EPDGTIKRLIYATSSLSGSGVSKRFSRSGDYSLTISSEDFVAYYCLOYASSPPTF 118
DB 60 KSGTSPKRWIYDTFKLTSGVDPDRFSGSGGTSLTISNWEAEDVATYYCCQWSRNPPTF 119

QY 119 GGTGKLEIKRADAAPTVSIFPPSSK 143
DB 120 GVGTKLEIKRADAAPTVSIFPPSS 144

RESULT 10
Q9JL84 PRELIMINARY; PRT; 107 AA.
ID Q9JL84
AC Q9JL84;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anti-mycosin immunoglobulin light chain variable region (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=20448942; PubMed=10992488;
RA Walkiel S., Liao L., Cunningham M.W., Diamond B.;
RT "T-Cell-dependent antibody response to the dominant epitope of streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive with cardiac myosin.";
RL Infect. Immun. 68:5803-5808(2000).
DR EMBL; AF206022; AAF69320.1; -
DR HSSP; P80362; 1WTL.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR NON_TER 1
FT NON_TER 107
SQ SEQUENCE 107 AA; 11648 MW; ACF9B1253ACAL1E5D CRC64;

Query Match 51.4%; Score 380; DB 11; Length 107;
Best Local Similarity 71.0%; Pred. No. 2.6e-32;
Matches 76; Conservative 5; Mismatches 26; Indels 0; Gaps 0;

QY 21 DIQMTQSPSSLSASLGQVSLTCTCRASQDIGNLHWLQEPDGTIKRLIYATSSLSGSPVK 80
DB 1 DIQMTQSTSLASLGRVAXXCSASQGISNXXWFQKQKPDGTVKLLIYTTSSLXSGVPS 60

QY 81 RFSGSRSGDYSLTISSEDFVAYYCLOYASSPPTFGGTGKLEIK 127
DB 61 RFSGSGGXDYSLTISNLEPDATYYCQISKFPPTFGGTGKLEIK 107

RESULT 11
Q99M37 PRELIMINARY; PRT; 238 AA.
ID Q99M37
AC Q99M37;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 26.3 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC002035; AAH02035.1; -  
DR HSP; P01679; 2FBJ  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_V.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS0835; IG\_LIKE; 2.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 238 AA; 26344 MW; FB2B06A0B801330A CRC64;

Query Match 50.5%; Score 374; DB 11; Length 238;  
Best Local Similarity 53.4%; Pred. No. 3.2e-31;  
Matches 79; Conservative 20; Mismatches 43; Indels 6; Gaps 2;  
QY 1 MGAPAIQLGILLFPFTRCDIOMTQSPSSLSASLGQVSLTCRASQDI-----GINLHW 55  
Db 1 MKLPVRL-LVLMFWIPASSSDVVMVTQPLSLPVLGDAQASISCRSSQSIHVSNGNTYLEW 59  
QY 56 LQEPDGTIKRLIYATSSLGSGVPRFSGSGSDYSLTSSLESDFVAYYCLQYASSP 115  
Db 60 YLQKPGQSKLLIYKVSNNRFGVPRDFSGSGGTDTFLKISRVEADLGVYFCQSHVP 119  
QY 116 YTFGGTGKLEIKRAADAAPTIVSIFPPSSK 143  
Db 120 YTFGGTGKLEIKRAADAAPTIVSIFPPSSSE 147

RESULT 12  
Q8TCD0 PRELIMINARY; PRT; 239 AA.  
ID Q8TCD0;  
AC Q8TCD0;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022362; AAH22362.1; -  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_V.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS0835; IG\_LIKE; 2.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 239 AA; 26234 MW; FACEDC3A3B03871D CRC64;

Query Match 50.2%; Score 371.5; DB 4; Length 239;  
Best Local Similarity 52.0%; Pred. No. 6e-31;  
Matches 77; Conservative 23; Mismatches 43; Indels 5; Gaps 1;  
QY 1 MGAPAIQLGILLFPFTRCDIOMTQSPSSLSASLGQVSLTCRASQ-----DIGINLHW 55  
Db 1 MRPLPAOLLGULLMFWPSSGSDVVMVTQPLSLPVLGQPASISCRSTQSLVSDGNTYLEW 60  
QY 56 LQEPDGTIKRLIYATSSLGSGVPRFSGSGSDYSLTSSLESDFVAYYCLQYASSP 115  
Db 61 FQORPGQSPRLIYKVSNNRFGVPRDFSGSGGTDTFLKISRVEADLGVYFCMQGTHWP 120  
QY 116 YTFGGTGKLEIKRAADAAPTIVSIFPPSSK 143

Db 121 STFGGTGKLEIKRTVAAPSVFIPPPSDE 148

RESULT 13  
Q8VCI6 PRELIMINARY; PRT; 238 AA.  
ID Q8VCI6;  
AC Q8VCI6;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical 26.2 kDa protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RA Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC019760; AAH19760.1; -  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_V.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 238 AA; 26224 MW; 35EC08E3DE5414AD CRC64;

Query Match 50.0%; Score 370; DB 11; Length 238;  
Best Local Similarity 52.7%; Pred. No. 8.5e-31;  
Matches 78; Conservative 21; Mismatches 43; Indels 6; Gaps 2;  
QY 1 MGAPAIQLGILLFPFTRCDIOMTQSPSSLSASLGQVSLTCRASQDI-----GINLHW 55  
Db 1 MKLPVRL-LVLMFWIPASSSDVVMVTQPLSLPVLGDAQASISCRSSQSIHVSNGNTYLEW 59  
QY 56 LQEPDGTIKRLIYATSSLGSGVPRFSGSGSDYSLTSSLESDFVAYYCLQYASSP 115  
Db 60 YLQKPGQSKLLIYKVSNNRFGVPRDFSGSGGTDTFLKISRVEADLGVYFCQSHVP 119  
QY 116 YTFGGTGKLEIKRAADAAPTIVSIFPPSSK 143  
Db 120 YTFGGTGKLEIKRAADAAPTIVSIFPPSSSE 147

RESULT 14  
Q8VDD0 PRELIMINARY; PRT; 134 AA.  
ID Q8VDD0;  
AC Q8VDD0;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Anti-MOG 212 variable light chain (Fragment).  
GN ANTI-MOG KAPPA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c;  
RA Chernajovsky Y.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c;  
RA Sembli P.;  
RT Targeting T cells to the CNS.  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ416331; CAC94866.1; -  
DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IG\_LIKE; 1.  
FT NON\_TER 134 134  
SQ SEQUENCE 134 AA; 14525 MW; CDFD9E2236E2D0CF CRC64;

Query Match 49.1%; Score 363.5; DB 11; Length 134;  
Best Local Similarity 57.8%; Pred. No. 1.9e-30;  
Matches 78; Conservative 17; Mismatches 37; Indels 3; Gaps 2;  
Qy 1 MGAPAQILGFLLLFPG--TRCDIQMTQSPSSLSASLGQRVSLTCRASQDIGINLHWLQ 58  
Db 1 MDFQVQIFSELLISASVILSRGQIVLTQSPAINSPGKVTMTCSASSISY-MHWYQ 59  
Qy 59 EPDGTIKRLIYATSSLGSGVPKRFSGSGRSDYSLTISSEDFVAYYCLOYASSPYTF 118  
Db 60 KPGTSPKRWIYDTSKLASGVPARFSGSGTSYSLTISSEADDAATYCHQRSSYPWF 119  
Qy 119 GGGTKLEIKRADAAP 133  
Db 120 GGGTKLEIKRADAAP 134

RESULT 15  
Q8K1F1 PRELIMINARY; PRT; 114 AA.  
AC O8K1F1;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Anti-VIPase light chain variable region (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALE/C; TISSUE=Hyperimmunized spleen;  
RA Zhou X.-X., Taguchi H., Planque S., Karle S., Nishiyama Y., Paul S.;  
RT "Innate proteolytic antibodies: Failed D-VIPase response to the D-  
FT entanionomer of VIP and identification of L-VIPase VL domains."  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF516284; AAM64202.1; -.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR07110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00409; IG; 1.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IG\_LIKE; 1.  
FT NON\_TER 1 114  
SQ SEQUENCE 114 AA; 12162 MW; 8BD9833DBF3EEFD1 CRC64;

Query Match 49.7%; Score 360.5; DB 11; Length 114;  
Best Local Similarity 63.2%; Pred. No. 3.2e-30;  
Matches 72; Conservative 16; Mismatches 25; Indels 1; Gaps 1;  
Qy 21 DIQMTQSPSSLSASLGQRVSLTCRASQDIGIN-LHWLQRPDGTIKRLIYATSSLGSGVP 79  
Db 1 DIVLTQSPAINSPGKVTMTCRASSVSSSYLHWYQKSGASPKLWIYSTSNLASGVP 60  
Qy 80 KRFGSGRSGDYSLTISSEDFVAYYCLOYASSPYTFGGTKLEIKRADAAP 133  
Db 61 ARFSGSGSGSYSLTISSEADDAATYCCQYHSYPTFFGGTKLEIKRADAAP 114



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 18:46:00 ; Search time 2089.42 Seconds  
(without alignments)  
9026.108 Million cell updates/sec

Title: US-08-836-455-3

Perfect score: 461

Sequence: 1 ATGAATGACAGCTGGTCTT.....CTGGTCCCTGGAAGCTTGGG 461

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*\*

1: gb\_ba:\*\*

2: gb\_htg:\*\*

3: gb\_in:\*\*

4: gb\_ov:\*\*

5: gb\_pat:\*\*

6: gb\_ph:\*\*

7: gb\_pl:\*\*

8: gb\_pr:\*\*

9: gb\_ro:\*\*

10: gb\_sts:\*\*

11: gb\_sy:\*\*

12: gb\_un:\*\*

13: gb\_vl:\*\*

14: gb\_vl:\*\*

15: em\_ba:\*\*

16: em\_fun:\*\*

17: em\_hum:\*\*

18: em\_in:\*\*

19: em\_mu:\*\*

20: em\_mu:\*\*

21: em\_or:\*\*

22: em\_ov:\*\*

23: em\_pat:\*\*

24: em\_ph:\*\*

25: em\_pl:\*\*

26: em\_ro:\*\*

27: em\_sts:\*\*

28: em\_un:\*\*

29: em\_vl:\*\*

30: em\_vl:\*\*

31: em\_htg\_hum:\*\*

32: em\_htg\_inv:\*\*

33: em\_htg\_other:\*\*

34: em\_htg\_mus:\*\*

35: em\_htg\_pln:\*\*

36: em\_htg\_rod:\*\*

37: em\_htg\_man:\*\*

38: em\_htg\_vrt:\*\*

39: em\_sy:\*\*

40: em\_htgo\_hum:\*\*

41: em\_htgo\_other:\*\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	461	100.0	461	6	AR164506	AR164506 Sequence
2	461	100.0	461	6	BD085738	BD085738 Methods O
3	407	88.3	411	10	AF124720	AF124720 Mus muscu
4	96	20.8	294	10	AF303872	AF303872 Mus muscu
5	84	18.2	471	10	AB050080	AB050080 Mus muscu
6	83	18.0	735	12	SC0250760	AJ250760 Mus muscu
7	77	16.7	902	12	XU049832	U49832 Synthetic s
8	76	16.5	458	6	AR083801	AR083801 Sequence
9	76	16.5	458	6	AR198720	AR198720 Sequence
10	76	16.5	458	6	AR275320	AR275320 Sequence
11	66	14.3	484	10	MUSIGH4C11	M54977 Mus musculu
12	66	14.3	490	10	MUSIG4C11A	M24785 Mouse anti-
13	61	13.2	453	10	MMU24114	U24114 Mus musculu
14	60	13.0	279	10	MMIGGCVRG	225453 M.musculus
15	60	13.0	282	10	MMIGGCVRB	225443 M.musculus
16	60	13.0	285	10	MMIGGCVRG	225445 M.musculus
17	60	13.0	290	10	MMIGGCVRG	225451 M.musculus
18	60	13.0	294	10	AF303873	AF303873 Mus muscu
19	60	13.0	302	10	MMIGGCVRG	225457 M.musculus
20	60	13.0	303	10	MMIGGCVRD	225447 M.musculus
21	60	13.0	305	10	MMIGGCVRG	225449 M.musculus
22	60	13.0	306	10	AY171955	AY171955 Mus muscu
23	60	13.0	321	10	AY171952	AY171952 Mus muscu
24	60	13.0	352	10	MMMD50C	273339 M.musculus
25	60	13.0	355	10	MMMD47C	273342 M.musculus
26	60	13.0	356	10	AY247151	AY247151 Mus muscu
27	60	13.0	358	10	MMMD01C	273357 M.musculus
28	60	13.0	358	10	MMMD52C	273338 M.musculus
29	60	13.0	360	10	AY229957	AY229957 Mus muscu
30	60	13.0	360	12	AF277091	AF277091 Synthetic
31	60	13.0	387	10	AY058907	AY058907 Mus muscu
32	60	13.0	457	10	AY178830	AY178830 Mus muscu
33	60	13.0	765	6	AX057984	AX057984 Sequence
34	60	13.0	765	6	AX057985	AX057985 Sequence
35	60	13.0	959	10	AF025445	AF025445 Mus muscu
36	60	13.0	1239	6	AX057945	AX057945 Sequence
37	60	13.0	1280	6	AX057947	AX057947 Sequence
38	59	12.8	336	10	AB090857	AB090857 Mus muscu
39	59	12.8	360	6	I73511	I73511 Sequence 14
40	59	12.8	360	10	AB058183	AB058183 Mus muscu
41	59	12.8	360	10	AB058198	AB058198 Mus muscu
42	59	12.8	363	10	AB057826	AB057826 Mus muscu
43	59	12.8	363	10	AB057827	AB057827 Mus muscu
44	59	12.8	363	10	AB090858	AB090858 Mus muscu
45	59	12.8	368	10	MMMD24	X97535 M.musculus

ALIGNMENTS

RESULT 1	AR164506	AR164506	Sequence 3	461 bp	DNA	linear	PAT 17-OCT-2001
LOCUS	AR164506	Sequence 3	from patent US 6274143.				
DEFINITION	AR164506	Sequence 3	from patent US 6274143.				
ACCESSION	AR164506	Sequence 3	from patent US 6274143.				
VERSION	AR164506.1	GI:16237556					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 461)						
AUTHORS	Chatterjee, M. and Foon, K. A.						
TITLE	Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11b10						
JOURNAL	Patent: US 6274143-A 3 14-AUG-2001;						

[illegible]



```

FEATURES             source            Location/Qualifiers
     1..471
     /organism="Mus musculus"
     /mol_type="mrna"
     /strain="BALB/c"
     /db_xref="taxon:10090"
     /cell_line="hybridoma 9H8"
     1..471
     /gene="VH9H8"
     11..>469
     /gene="VH9H8"
     /codon_start=1
     /product="anti-A/u antibody"
     /protein_id="BAB87192.1"
     /db_xref="GI:19909936"
     /translation="MGWTGVFLILLITVGHQAYLQSGAELVRSGASVKMSCKRAS
GYNFTSNMHWKOTPGQGLEWIGYIFLNAGTNYNOKFKGKATLTADTSSSTAYMQLS
SLRSEDSAVFCARGEYGNFWFAYWGQGLTVTSAAKTTAQPVPYPLAPGSL"
BASE COUNT          114 a 126 c 121 g 110 t
ORIGIN

Query Match          18.2%; Score 84; DB 10; Length 471;
Best Local Similarity 100.0%; Pred. No. 6.4e-37;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 GCAGACATCTCTCCAGCACAGCCTACATGCAGATCAGACGCTTGACATCTGAAGACTCT 330
      |||||||
DB 278 GCAGACATCTCTCCAGCACAGCCTACATGCAGATCAGACGCTTGACATCTGAAGACTCT 337
      |||||||

QY 331 CGGTCATTCTCTGCAAGAGG 354
      |||||||
DB 338 CGGTCATTCTCTGCAAGAGG 361
      |||||||

RESULT 6
SC0250760
LOCUS              735 bp mRNA linear SYN 11-MAY-2000
DEFINITION         Mus musculus synthetic construct for anti-guinea pig C5 ScFv
                    antibody, clone E10.
VERSION            AJ250760
KEYWORDS           antibody; heavy chain; immunoglobulin superfamily; light chain;
                    ScFv; variable region.
SOURCE             synthetic construct
ORGANISM           artificial sequences.
REFERENCE          1
AUTHORS            Link,C., Hawlisch,H., Meyer zu Vilsendorf,A., Gylteruez,S., Nagel,E.
                    and Koehl,J.
TITLE              Selection of phage-displayed anti-guinea pig C5 or C5a antibodies
                    and their application in xenotransplantation
JOURNAL            Mol. Immunol. 38, 1235-1247 (1999)
AUTHORS            Link,C.
TITLE              Direct Submission
JOURNAL            Submitted (01-NOV-1999) Link C., Medizinische Mikrobiologie,
                    Medizinische Hochschule Hannover, Carl-Neuberg-Str. 1, 30625
                    Hannover, GERMANY
FEATURES           Location/Qualifiers
     1..735
     /organism="synthetic construct"
     /mol_type="mrna"
     /db_xref="taxon:32630"
     /clone="E10"
     /focus
     1..366
     /organism="Mus musculus"
     /mol_type="mrna"
     /db_xref="taxon:10090"
     412..735
     /organism="Mus musculus"
     /mol_type="mrna"
     /db_xref="taxon:10090"

FEATURES           source            Location/Qualifiers
     1..>735
     /note="synthetic"
     /codon_start=1
     /product="anti-guinea pig C5 ScFv antibody"
     /protein_id="CAB60133.1"
     /db_xref="GI:6272273"
     /translation="EVLQOQSGAELVRSGASVKMSCKASGYTFARYNVHWKOTPGQG
LEWIGYTPRDGNTNOKFKGKATLTADTSSSTAYMQLSLSLSEDSAVFCRRSRDA
YVVGAMDFWQGTSTVTSGGGGGGGGSDIVMTQSQKFMSTSVGDRVSVTC
KASQNYCTNVAWYQKPGQSPKALIVSASRYSGVPDRFTGSGGTDFTLTISNQVE
DLAEYFCQYNSIPYTFGGTNLEKR"
V_region           1..366
                    /product="immunoglobulin heavy chain variable region"
misc_feature       367..411
                    /note="(Gly4Ser)3 linker sequence"
V_region           412..735
                    /product="immunoglobulin light chain variable region"
BASE COUNT          194 a 170 c 206 g 165 t
ORIGIN

Query Match          18.0%; Score 83; DB 12; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.4e-36;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 CATTGACTGCAGACATCTCTCCAGCACAGCCTACATGCAGATCAGACGCTGCATCTG 322
      |||||||
DB 206 CATTGACTGCAGACATCTCTCCAGCACAGCCTACATGCAGATCAGACGCTGCATCTG 265
      |||||||

QY 323 AAGACTCTGGGCTATTTCGT 345
      |||||||
DB 266 AAGACTCTGGGCTATTTCGT 288
      |||||||

RESULT 7
XXU49832
LOCUS              902 bp mRNA linear SYN 24-MAR-1996
DEFINITION         Synthetic single chain Fv antibody against potato virus V coat
                    protein, mRNA, partial cds.
VERSION            U49832
KEYWORDS           U49832.1 GI:1236090
SOURCE             synthetic construct
ORGANISM           synthetic construct
                    artificial sequences.
REFERENCE          1 (bases 1 to 902)
AUTHORS            Chen,Z.C., Cockburn,W., Torrance,L., Barker,H. and Whitelam,G.C.
                    Cytoplasmic accumulation of a soluble functional scFv protein to a
                    plant virus expressed as a thiorodoxin fusion in Escherichia coli
TITLE              Unpublished
JOURNAL            Submitted (23-FEB-1996) Z C. Chen, Botany, Univ. of Leicester,
                    University Road, Leicester LE1 7RH, UK
FEATURES           Location/Qualifiers
     1..902
     /organism="synthetic construct"
     /mol_type="mrna"
     /db_xref="taxon:32630"
     /note="PCR synthesized gene from mouse hybridoma cell line
                    SCR 40"
     <1..885
     /codon_start=1
     /product="single chain Fv antibody"
     /protein_id="AA92722.1"
     /db_xref="GI:1236091"
     /translation="SRPMADVQIIQTASLSASVGEVTVTTCRASEHIYSYLAWYQOK
QKSPQLLVYSAKTLAEQVPSRFSGSGGQFSLKINSIQLPEDFGYSYQHHYDTPRT
FGGKTKLEIRADAAPTVSIFPPSPGGGGGGGGGGSDIVMTQSQKFMSTSVGDRVSVTC
SVKMSCKASGYTFSTYNMHWKOTPGQLEWIGYIYPNGGTIYNQKFKGKATLTADT
SSSTANNQISSLSSEDSAVFCARGDYRNDPFDPMWGGTTLTVSSAKTTPPSVYPLAP

FEATURES           source            Location/Qualifiers
     1..735
     /organism="synthetic construct"
     /mol_type="mrna"
     /db_xref="taxon:32630"
     /clone="E10"
     /focus
     1..366
     /organism="Mus musculus"
     /mol_type="mrna"
     /db_xref="taxon:10090"
     412..735
     /organism="Mus musculus"
     /mol_type="mrna"
     /db_xref="taxon:10090"

```

```
BASE COUNT      223 a      239 c      234 g      206 t
ORIGIN
LEGAGSGOPELAPEDPED"
Query Match      16.7%; Score 77; DB 12; Length 902;
Best Local Similarity 99.2%; Pred. No. 7.2e-33;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 77 CTGGGGCTGAGCTGTGGGTCTGGGGCTCAGTGAAGATGTCCTGCAAGGCTTCTGGCT 136
Db 452 CTGGGGCTGAGCTGTGGGTCTGGGGCTCAGTGAAGATGTCCTGCAAGGCTTCTGGCT 511
QY 137 ACACATTGACCACTTACATATGCACTGGTGAACACACACCTGGACAGGCGCTGGAAAT 196
Db 512 ACACATTTACCACTTACATATGCACTGGTGAACACACACCTGGACAGGCGCTGGAAAT 571
QY 197 GGATTGGA 204
Db 572 GGATTGGA 579
RESULT 8
AR083801
LOCUS      AR083801      458 bp      DNA      linear      PAT 01-SEP-2000
DEFINITION Sequence 3 from patent US 5977316.
ACCESSION AR083801
VERSION AR083801.1 GI:10010572
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 458)
AUTHORS Chatterjee,M., Foon,K.A. and Chatterjee,S.K.
TITLE Monoclonal antibody 1A7 and related polypeptides
JOURNAL Patent: US 5977316-A 3 02-NOV-1999;
FEATURES Location/Qualifiers
source 1..458
/organism="unknown"
BASE COUNT      106 a      131 c      114 g      107 t
ORIGIN
Query Match      16.5%; Score 76; DB 6; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.9e-32;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 366 TGCTCTGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 425
Db 363 TGCTCTGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 422
QY 426 CCCACCCGCTCTATCCA 441
Db 423 CCCACCCGCTCTATCCA 438
RESULT 9
AR198720
LOCUS      AR198720      458 bp      DNA      linear      PAT 20-APR-2002
DEFINITION Sequence 3 from patent US 6355244.
ACCESSION AR198720
VERSION AR198720.1 GI:20248794
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 458)
AUTHORS Foon,K.A. and Chatterjee,M.
TITLE Methods and compositions for the treatment of psoriasis
JOURNAL Patent: US 6355244-A 3 12-MAR-2002;
FEATURES Location/Qualifiers
source 1..458
/organism="unknown"
BASE COUNT      106 a      131 c      114 g      107 t
ORIGIN
LEGAGSGOPELAPEDPED"
Query Match      16.7%; Score 77; DB 12; Length 902;
Best Local Similarity 99.2%; Pred. No. 7.2e-33;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 77 CTGGGGCTGAGCTGTGGGTCTGGGGCTCAGTGAAGATGTCCTGCAAGGCTTCTGGCT 136
Db 452 CTGGGGCTGAGCTGTGGGTCTGGGGCTCAGTGAAGATGTCCTGCAAGGCTTCTGGCT 511
QY 137 ACACATTGACCACTTACATATGCACTGGTGAACACACACCTGGACAGGCGCTGGAAAT 196
Db 512 ACACATTTACCACTTACATATGCACTGGTGAACACACACCTGGACAGGCGCTGGAAAT 571
QY 197 GGATTGGA 204
Db 572 GGATTGGA 579
RESULT 8
AR083801
LOCUS      AR083801      458 bp      DNA      linear      PAT 01-SEP-2000
DEFINITION Sequence 3 from patent US 5977316.
ACCESSION AR083801
VERSION AR083801.1 GI:10010572
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 458)
AUTHORS Chatterjee,M., Foon,K.A. and Chatterjee,S.K.
TITLE Monoclonal antibody 1A7 and related polypeptides
JOURNAL Patent: US 5977316-A 3 02-NOV-1999;
FEATURES Location/Qualifiers
source 1..458
/organism="unknown"
BASE COUNT      106 a      131 c      114 g      107 t
ORIGIN
Query Match      16.5%; Score 76; DB 6; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.9e-32;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 366 TGCTCTGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 425
Db 363 TGCTCTGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 422
QY 426 CCCACCCGCTCTATCCA 441
Db 423 CCCACCCGCTCTATCCA 438
RESULT 9
AR198720
LOCUS      AR198720      458 bp      DNA      linear      PAT 20-APR-2002
DEFINITION Sequence 3 from patent US 6355244.
ACCESSION AR198720
VERSION AR198720.1 GI:20248794
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 458)
AUTHORS Foon,K.A. and Chatterjee,M.
TITLE Methods and compositions for the treatment of psoriasis
JOURNAL Patent: US 6355244-A 3 12-MAR-2002;
FEATURES Location/Qualifiers
source 1..458
/organism="unknown"
BASE COUNT      106 a      131 c      114 g      107 t
ORIGIN
LEGAGSGOPELAPEDPED"
Query Match      16.5%; Score 76; DB 6; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.9e-32;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 366 TGCTCTGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 425
Db 363 TGCTCTGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 422
QY 426 CCCACCCGCTCTATCCA 441
Db 423 CCCACCCGCTCTATCCA 438
RESULT 11
MUSIGH4C11
LOCUS      MUSIGH4C11      484 bp      mRNA      linear      ROD 12-JUN-2000
DEFINITION partial cds.
ACCESSION M54977
VERSION M54977.1 GI:194528
KEYWORDS C-region; V-region; immunoglobulin heavy chain.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 484)
AUTHORS Sood,A.K., Cheng,H.L. and Kohler,H.
TITLE An efficient and general method for sequencing immunoglobulin mRNAs
JOURNAL J. Immunol. Methods 95 (2), 227-235 (1986)
MEDLINE 87084812
PUBMED 2432131
COMMENT Original source text: Mouse anti-idiotypic hybridoma cell line 4C11, cDNA to mRNA.
FEATURES Location/Qualifiers
source 1..484
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
```

```

/map="chromosome 12"
/cell_line="4C11"
/tissue_type="anti-idiotypic hybridoma"
1. .484
/gene="Igh"
CDS 37. .>484
/gene="Igh"
/codon_start=1
/product="immunoglobulin heavy chain precursor"
/protein_id="AAA51051.1"
/db_xref="GI:553945"
/translation="MWSWIFLELLSSTAGVLSEVLOQSGPELLPAGSVKISCKAS
GYFTDITHHWKQSGKSLGWIGGDPNDYTFYNEKFKATLTVDKSSITAYMEL
RSTSDSAVYICASYDYGALDYGQGSTVTSAAKTPPSVYPLAP"
37. .93
/gene="Igh"
mat_peptide 94. .>484
/gene="Igh"
/product="immunoglobulin heavy chain"
445. .>484
/gene="Igh"
C_region 120 a 134 c 112 g 118 t
BASE COUNT 120 a 134 c 112 g 118 t
ORIGIN
Query Match 14.3%; Score 66; DB 10; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.9e-26;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 GGTGCTGCTGACTACTGGGCTCAAGGAACCTCAGTCACCGTCTCCTCAGCAAAAGACA 423
|||||
Db 397 GGTGCTGCTGACTACTGGGCTCAAGGAACCTCAGTCACCGTCTCCTCAGCAAAAGACA 456

QY 424 CCCCCA 429
|||||
Db 457 CCCCCA 462

RESULT 12
MUSIG4C11A 490 bp mRNA linear ROD 27-APR-1993
LOCUS Mouse anti-idiotypic immunoglobulin heavy chain variable, constant,
DEFINITION and complementarity determining regions 1-3 (4C11) mRNA.
ACCESSION M24785
VERSION M24785.1 GI:194133
KEYWORDS C-region; V-region; anti-idiotypic antibody; complementarity
determining region; immunoglobulin heavy chain.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 490)
AUTHORS Cheng,H.L., Sood,A.K., Ward,R.E., Kieber-Emmons,T. and Kohler,H.
TITLE Structural basis of stimulatory anti-idiotypic antibodies
JOURNAL Mol. Immunol. 25 (1), 33-40 (1988)
MEDLINE 86142863
PUBMED 3125424
COMMENT Original source text: Mus musculus (strain BALB/c; sub_species
domesticus) cDNA to mRNA.
FEATURES
source 1. .490
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
gene 1. .490
/gene="4C11"
V_region 1. .441
/gene="4C11"
gene 184. .198
/gene="CDR1"
misc_feature 184. .198
/gene="CDR1"
/gene="CDR1"

```

```

/note="complementarity determining region"
241. .291
/gene="CDR2"
misc_feature 241. .291
/gene="CDR2"
/note="complementarity determining region"
388. .411
/gene="CDR3"
misc_feature 388. .411
/gene="CDR3"
/note="complementarity determining region"
442. .490
/gene="4C11"
C_region 120 a 134 c 116 g 120 t
BASE COUNT 120 a 134 c 116 g 120 t
ORIGIN
Query Match 14.3%; Score 66; DB 10; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.9e-26;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 GGTGCTGCTGACTACTGGGCTCAAGGAACCTCAGTCACCGTCTCCTCAGCAAAAGACA 423
|||||
Db 397 GGTGCTGCTGACTACTGGGCTCAAGGAACCTCAGTCACCGTCTCCTCAGCAAAAGACA 456

QY 424 CCCCCA 429
|||||
Db 457 CCCCCA 462

RESULT 13
MMU24114 453 bp mRNA linear ROD 20-APR-1995
LOCUS Mus musculus immunoglobulin F9.13.7 heavy chain mRNA, partial cds.
DEFINITION U24114
ACCESSION U24114
VERSION U24114.1 GI:777410
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 453)
AUTHORS Lescar,J., Pellegriani,M., Souchon,H., Tello,D., Poljak,R.J.,
Peterson,N.C., Greene,M.I. and Alzari,P.M.
TITLE Crystal structure of a cross-reaction complex between Fab F9.13.7
and Guinea-fowl lysozyme
JOURNAL J. Biol. Chem. (1995) In press
REFERENCE 2 (bases 1 to 453)
AUTHORS Peterson,N.C.
TITLE Direct Submission
JOURNAL Submitted (03-APR-1995) Norman C. Peterson, Pathology and
Laboratory Medicine, University of Pennsylvania School of Medicine,
252 John Morgan Bldg., 36th St. and Hamilton Walk, Philadelphia, PA
19104, USA
FEATURES
source 1. .453
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/cell_type="lymphocyte/hybridoma"
<1. .>453
/codon_start=1
/product="immunoglobulin F9.13.7 heavy chain"
/protein_id="AAA65180.1"
/db_xref="GI:777411"
/translation="AELVPGASVKLSCKASGYTFYVNMHWKQSGOGLEWIGELD
PDSYPNIEKFKATLTVDKSSITAYMQLSDTSASVYICASYIGTSGYGLD
YWGQGSTVTSAAKTPPSVYPLAPGSAAGTNSWTLGCLVKGYPPEPV"
64. .81
/misc_feature 64. .81
label="CDR1"
misc_feature 124. .174
label="CDR2"
misc_feature 268. .309
label="CDR2"

```

```

misc_feature      /label=CDR3
343
/label=begin_CH1
BASE COUNT      107 a 127 c 112 g 107 t
ORIGIN
Query Match      13.0%; Score 61; DB 10; Length 453;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 369 TCTGGACTACTGGGCTCAAGGACCTCAGTCACCGTCTCTCAGCCAAACGACACCCCC 428
Db 300 TCTGGACTACTGGGCTCAAGGACCTCAGTCACCGTCTCTCAGCCAAACGACACCCCC 359

Qy 429 A 429
Db 360 A 360

RESULT 14
MMIGGCVRG      279 bp mRNA linear ROD 13-OCT-1993
LOCUS      M.musculus immunoglobulin gamma heavy chain (DBA/1) gene, v region.
DEFINITION
ACCESSION      Z25453
VERSION      Z25453.1 GI:407822
KEYWORDS      heavy chain; IgG gene; immunoglobulin; variable region.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 279)
AUTHORS      Mo,J.A., Bona,C.A. and Holmdahl,R.
TITLE      Variable region gene selection of immunoglobulin G expressing B
cells with specificity for a defined epitope on type II collagen
JOURNAL      Eur. J. Immunol. (1993) In press
REFERENCE      2 (bases 1 to 279)
AUTHORS      Mo,J.A., Bona,C.A. and Holmdahl,R.
TITLE      Variable region gene selection of immunoglobulin G-expressing B
cells with specificity for a defined epitope on type II collagen
JOURNAL      Eur. J. Immunol. 23 (10), 2503-2510 (1993)
MEDLINE      94009207
PUBMED      7691608
REFERENCE      3 (bases 1 to 279)
AUTHORS      Mo,J.A.
TITLE      Direct Submission
JOURNAL      Submitted (02-AUG-1993) John A Mo, Department of Medical and
Physiological, Department of, Medical and Physiological Chemistry,
Husargatan 3, Uppsala, 75123, Sweden
FEATURES      Location/Qualifiers
source      1..279
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="DBA/1"
             /db_xref="taxon:10090"
             /chromosome="12"
             /cell_line="CIICB340 hybridoma"
             /cell_type="B cell hybridoma"
             /tissue_type="lymph node"
             /dev_stage="adult"
             1..279
             /gene="IgG"
             <1..>279
             /gene="IgG"
             /codon_start=1
             /product="immunoglobulin gamma heavy chain"
             /protein_id="CAA80940.1"
             /db_xref="GI:407823"
             /translation="GAELVPGASVKMSCKASGYFTSYNMHWKQTPGQGLEWIGAI
             YPNGDTSYNQKFKGKATLTADKSSSTAYMQLSLTSEDSAVYYCARRE"
             1..279
             /gene="IgG"
             /dev_stage="adult"
             76 a 66 c 77 g 60 t

V_region
BASE COUNT      76 a 66 c 77 g 60 t
ORIGIN
Query Match      13.0%; Score 60; DB 10; Length 282;
Best Local Similarity 100.0%; Pred. No. 6.1e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 ACCAGTTACAATATGCACTGGGTAAAGCAGACACCTGGACAGGCGCTGGAATGGATTGGA 204
Db 67 ACCAGTTACAATATGCACTGGGTAAAGCAGACACCTGGACAGGCGCTGGAATGGATTGGA 126

RESULT 15
MMIGGCVRG      282 bp mRNA linear ROD 13-OCT-1993
LOCUS      M.musculus immunoglobulin gamma heavy chain (DBA/1) gene, v region.
DEFINITION
ACCESSION      Z25443
VERSION      Z25443.1 GI:407812
KEYWORDS      heavy chain; IgG gene; immunoglobulin; variable region.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 282)
AUTHORS      Mo,J.A., Bona,C.A. and Holmdahl,R.
TITLE      Variable region gene selection of immunoglobulin G expressing B
cells with specificity for a defined epitope on type II collagen
JOURNAL      Eur. J. Immunol. (1993) In press
REFERENCE      2 (bases 1 to 282)
AUTHORS      Mo,J.A., Bona,C.A. and Holmdahl,R.
TITLE      Variable region gene selection of immunoglobulin G-expressing B
cells with specificity for a defined epitope on type II collagen
JOURNAL      Eur. J. Immunol. 23 (10), 2503-2510 (1993)
MEDLINE      94009207
PUBMED      7691608
REFERENCE      3 (bases 1 to 282)
AUTHORS      Mo,J.A.
TITLE      Direct Submission
JOURNAL      Submitted (02-AUG-1993) John A Mo, Department of Medical and
Physiological, Department of, Medical and Physiological Chemistry,
Husargatan 3, Uppsala, 75123, Sweden
FEATURES      Location/Qualifiers
source      1..282
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="DBA/1"
             /db_xref="taxon:10090"
             /chromosome="12"
             /cell_line="CIICB20 hybridoma"
             /cell_type="B cell hybridoma"
             /tissue_type="lymph node"
             /dev_stage="adult"
             1..282
             /gene="IgG"
             <1..>282
             /gene="IgG"
             /codon_start=1
             /product="immunoglobulin gamma heavy chain"
             /protein_id="CAA80930.1"
             /db_xref="GI:407813"
             /translation="GAELVPGASVKMSCKASGYFTSYNMHWKQTPGQGLEWIGAI
             YPNGDTSYNQKFKGKATLTADKSSSTAYMQLSLTSEDSAVYYCARGEA"
             1..282
             /gene="IgG"
             74 a 80 g 60 t

V_region
BASE COUNT      74 a 80 g 60 t
ORIGIN
Query Match      13.0%; Score 60; DB 10; Length 282;
Best Local Similarity 100.0%; Pred. No. 6.1e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 ACCAGTTACAATATGCACTGGGTAAAGCAGACACCTGGACAGGCGCTGGAATGGATTGGA 204
Db 67 ACCAGTTACAATATGCACTGGGTAAAGCAGACACCTGGACAGGCGCTGGAATGGATTGGA 126

```

Search completed: August 30, 2003, 21:04:56  
Job time : 2091.42 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 17:38:10 ; Search time 201.688 Seconds  
(without alignments)  
6170.144 Million cell updates/sec

Title: US-08-836-455-3  
Perfect score: 461  
Sequence: 1 ATGGAATGACAGCTGGGTCTT.....CTGGTCCCTGGAAGCTTGGG 461

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :			
N_Geneseq_19Jun03.*			
1:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*		
2:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*		
3:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*		
4:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*		
5:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*		
6:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*		
7:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*		
8:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*		
9:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*		
10:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*		
11:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*		
12:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*		
13:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*		
14:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*		
15:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*		
16:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*		
17:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*		
18:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*		
19:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*		
20:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*		
21:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*		
22:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*		
23:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*		
24:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*		
25:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	461	100.0	461	18	AAT85150 Murine monoclonal
2	461	100.0	461	20	AAV83773 Antibody 11D10 hea
3	461	100.0	461	25	AA151274 Mouse 11D10 antibo
4	76	16.5	458	17	AA131333 Anti-idiotype mono
5	76	16.5	458	20	AA131366 MAb 1A7 heavy chai
6	76	16.5	458	20	AA189553 Heavy chain variab
7	76	16.5	458	20	AA160630 Monoclonal antibod
8	67	14.5	861	16	AAQ81500 sfv anti-rev seque

9	67	14.5	861	18	AAT45347	Single chain sfv a
10	60	13.0	765	22	AAC86590	DNA encoding a fus
11	60	13.0	765	22	AAC86591	DNA encoding a fus
12	60	13.0	1239	22	AAC88563	DNA encoding a fus
13	60	13.0	1280	22	AAC88564	Anti-CD20 single c
14	60	13.0	1925	21	AA15019	DNA encoding a CD-
15	59	12.8	360	18	AAT96345	cDNA for Ig heavy
16	59	12.8	420	22	AAF81910	Anti-CAL25 bifunct
17	59	12.8	447	20	AA57786	Anti-HCV Ser/Thr p
18	59	12.8	458	10	AA91146	2H7 Vh sequence
19	59	12.8	459	20	AAV82357	Mouse antibody 2H7
20	59	12.8	459	22	AAH22070	2H7 heavy chain va
21	59	12.8	470	16	AAQ99892	Mouse VLA-4 antibo
22	59	12.8	470	18	AAT74760	Alpha-4 integrin m
23	59	12.8	477	18	AAT70806	Mouse anti-idiotyp
24	59	12.8	486	18	AAT59339	MH1 monoclonal ant
25	59	12.8	491	18	AAT70868	2H7 heavy chain va
26	59	12.8	491	19	AAV18557	Mouse 2H7 antibody
27	59	12.8	491	19	AAV03926	Mouse 2H7 antibody
28	59	12.8	491	19	AAV18593	Mouse 2H7 antibody
29	59	12.8	518	18	AAT36316	2H7 antibody heavy
30	59	12.8	520	18	AAT51042	Coding sequence fo
31	59	12.8	588	14	AAQ43385	H-chain V-region o
32	59	12.8	626	25	ACC44926	TSH receptor antib
33	59	12.8	626	25	ACC44927	TSH receptor antib
34	59	12.8	736	25	ABX16571	Mouse DNA encoding
35	59	12.8	1347	11	AAQ05708	Heavy chain of ant
36	59	12.8	1347	14	AAQ51534	Sequence encoding
37	59	12.8	1347	17	AAT23056	Murine anti-BGH MA
38	59	12.8	1347	17	AAT13733	Anti-BGH monoclon
39	59	12.8	1528	5	AA40024	Combined cDNA inse
40	59	12.8	1528	5	AA40025	mRNA encoding gamm
41	59	12.8	1553	16	AAQ79930	Anti-tobacco mosai
42	59	12.8	1797	17	AAT15733	3B1 single chain a
43	59	12.8	1848	18	AAT96346	Chimeric gene cont
44	59	12.8	3343	11	AAQ04655	Plasmid p10169 enc
45	57	12.4	399	16	AAQ90425	DNA encoding anti-

ALIGNMENTS

RESULT 1	
AAT85150	AAT85150 standard; cDNA; 461 BP.
ID	AAT85150 standard; cDNA; 461 BP.
XX	
AC	AAT85150;
XX	
DT	25-MAR-2003 (updated)
DT	04-JAN-1998 (first entry)
XX	
DE	Murine monoclonal anti-idiotype antibody 11D10 VH cDNA.
XX	
KW	Monoclonal antibody 11D10; anti-idiotype antibody; mucin;
KW	human milk fat globule; HMF; tumour; breast cancer; vaccine; ss.
XX	
OS	Mus musculus.
XX	
FH	Key
FT	sig_peptide
FT	mat_peptide
FT	Location/Qualifiers
FT	1..57
FT	/*tag= a
FT	58..461
FT	/*tag= b
XX	
PN	WO9722699-A2.
XX	
PD	26-JUN-1997.
XX	
PF	19-DEC-1996;
XX	96WO-US20757.
XX	
PR	20-DEC-1995;
PR	95US-0575762.
PR	26-JAN-1996;
PR	96US-0591965.
PR	13-DEC-1996;
PR	96US-0766350.





```

XX 17-JAN-1995; 95US-0372676.
PR 16-JAN-1996; 96US-0591196.
PR 17-JAN-1995; 95US-0372676.
PR 16-JAN-1996; 96US-0591196.
XX (KENT ) UNIV KENTUCKY.
XX Chatterjee M, Chatterjee SK, Foon KA;
XX WPI: 1996-354530/35.
DR P-PSDB; AAW03200.
XX Monoclonal antibody 1A7 and related polynucleotide(s) and
PT polypeptide(s) - useful to treat or palliate a GD2-associated
PT disease, e.g. melanoma and glioma
XX Claim 11; Fig 2; 141pp; English.
XX The present sequence encodes the murine anti-idiotypic monoclonal
CC antibody (MAB) 1A7 variable heavy chain. MAB 1A7 was raised against
CC the anti-ganglioside 2 (GD2) MAB 14G2a, which binds an unique
CC epitope of GD2. As the glycosphingolipid GD2 is expressed at high
CC density by human neuroectodermal tumours, e.g. malignant melanoma,
CC neuroblastoma, glioma, soft tissue sarcoma and small cell carcinoma
CC of the lung, MAB 1A7, or its cDNA can be used in a vaccine to treat
CC or palliate such diseases. They can also be used to reduce the
CC risk of recurrence of a clinically detectable tumour, and detect an
CC anti-GD2 Ab bound to a tumour cell.
CC MAB 1A7 overcomes immune tolerance and induces an immune response
CC against GD2, which comprises anti-GD2 Ab (humoral response) and
CC GD2-specific cells (cellular response). It can be used to purify
CC anti-1A7 (Ab3), anti-GD2 (Ab1') or 14G2a (Ab1), detect anti-1A7 or
CC anti-GD2 in a sample or measure the level of cellular anti-1A7 or
CC anti-GD2 activity.
CC The cDNA can be used in expression systems for 1A7 prodn., and in
CC the prepn. of probes and primers to respectively assay for 1A7
CC cDNA, and amplify desired polynucleotides for use in gene therapy.
CC (Updated on 25-MAR-2003 to correct PR field.)
XX Sequence 458 BP; 106 A; 131 C; 114 G; 107 T; 0 other;
SQ
Query Match 16.5%; Score 76; DB 17; Length 458;
Best Local Similarity 100.0%; Pred. No. 3.2e-25;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 366 TGCCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 425
DB 363 TGCCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 422
OY 426 CCCACCGCTCTATCCA 441
DB 423 CCCACCGCTCTATCCA 438
RESULT 5
AAZ31366
ID AAZ31366 standard; cDNA; 458 BP.
XX AAZ31366;
XX MAB 1A7 heavy chain variable region encoding cDNA.
XX 07-FEB-2000 (first entry)
XX heavy chain variable region; antibody 1A7; T cell response; melanoma;
KW neuroblastoma; glioma; soft tissue carcinoma; small cell carcinoma;
KW tumor-associated antigen; ss.
XX Synthetic.
OS Mus sp.
XX US5977316-A.
XX
XX 02-NOV-1999.
XX 16-JAN-1996; 96US-0591196.
XX 17-JAN-1995; 95US-0372676.
XX (KENT ) UNIV KENTUCKY.
XX Foon KA, Chatterjee SK, Chatterjee M;
XX WPI: 1999-619711/53.
DR P-PSDB; AAY49210.
XX Monoclonal antibody 1A7 which elicits an anti-GD2 immunological
PT response, useful for the development of products for the detection and
PT treatment of cancers
XX Disclosure; Fig 2; 74pp; English.
XX The invention provides a monoclonal antibody (MAB) designated 1A7, which
CC elicits an anti-GD2 (tumor-associated antigen) immunological response in
CC humans. MAB 1A7 has defined light and heavy chain variable region
CC sequences. The MAB 1A7 and polypeptides can be used for eliciting an
CC anti-GD2 immune response. The polypeptides can also be used for detecting
CC or purifying anti-GD2 antibody. The products can be used for treating GD2
CC -associated diseases, e.g. melanoma, neuroblastoma, glioma, soft tissue
CC carcinoma, and small cell carcinoma. They can be used for palliating the
CC disease or for reducing the risk of recurrence. The present sequence
CC represents the cDNA encoding the heavy chain variable region of MAB 1A7.
XX Sequence 458 BP; 106 A; 131 C; 114 G; 107 T; 0 other;
SQ
Query Match 16.5%; Score 76; DB 20; Length 458;
Best Local Similarity 100.0%; Pred. No. 3.2e-25;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 366 TGCCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 425
DB 363 TGCCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 422
OY 426 CCCACCGCTCTATCCA 441
DB 423 CCCACCGCTCTATCCA 438
RESULT 6
AAZ89553
ID AAZ89553 standard; cDNA; 458 BP.
XX AAZ89553;
XX 06-OCT-1999 (first entry)
XX Heavy chain variable region of MAB 1A7.
XX heavy chain variable region; antibody 1A7; T cell response; melanoma;
KW ganglioside GD2; CDR; complementarity determining region; carcinoma; ss.
XX Mus musculus.
XX Key Location/Qualifiers
FH CDS 1..456
FT /*tag= a
FT /*product= "Vh chain MAB 1A7"
FT /*note= "No stop codon given"
FT sig_peptide 1..57
FT /*tag= b
FT mat_peptide 58..456
FT /*tag= c
XX US5935821-A.

```

PD 10-AUG-1999.  
 XX  
 PF 21-NOV-1996; 9605-0752844.  
 XX  
 PR 21-NOV-1996; 9605-0752844.  
 PR 17-JAN-1995; 9505-0372676.  
 PR 16-JAN-1996; 9605-0591196.  
 XX  
 PA (KENT ) UNIV KENTUCKY.  
 XX  
 PI Chatterjee M, Chatterjee SK, Foon KA;  
 XX  
 DR WPI; 1999-457600/38.  
 DR P-PSDB; AAY28469.  
 XX  
 PT Anti-GD2 immunological peptides useful for the treatment of tumours  
 XX especially melanomas and small cell carcinomas  
 XX  
 PS Claim 7; Fig 2; 84pp; English.  
 CC The sequence is the variable heavy chain region of monoclonal  
 CC anti-idiotypic antibody 1A7. The polypeptide encoded by this sequence has  
 CC three CDRs (complementarity determining regions). When administered to  
 CC an individual the 1A7 antibody induces an immune response against  
 CC ganglioside GD2. The light chain variable region of the 1A7 antibody  
 CC (AAY28468) is also capable of eliciting an anti GD2 response in mammals.  
 CC Both the heavy and light chain variable regions of the 1A7 antibody  
 CC produce anti-GD2 T cell and antibody responses. The peptides and  
 CC antibodies may be useful for the modulation of ganglioside GD2, and  
 CC particularly for the treatment of GD2-associated tumours (e.g. melanoma,  
 CC neuroblastoma, glioma, soft tissue sarcoma, and small cell carcinoma  
 CC (including small cell lung cancer)).  
 XX  
 SQ Sequence 458 BP; 106 A; 131 C; 114 G; 107 T; 0 other;  
 Query Match 16.5%; Score 76; DB 20; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-25;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 366 TGCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCCAAACGACACC 425  
 DB 363 TGCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCCAAACGACACC 422  
 QY 426 CCCACCCGCTCTATCCA 441  
 DB 423 CCCACCCGCTCTATCCA 438  
 Query Match 16.5%; Score 76; DB 20; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-25;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 366 TGCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCCAAACGACACC 425  
 DB 363 TGCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCCAAACGACACC 422  
 QY 426 CCCACCCGCTCTATCCA 441  
 DB 423 CCCACCCGCTCTATCCA 438  
 RESULT 7  
 AAX60630  
 ID AAX60630 standard; cDNA; 458 BP.  
 XX  
 AC AAX60630;  
 XX  
 DT 03-AUG-1999 (first entry)  
 XX  
 DE Monoclonal antibody 1A7 heavy chain variable region encoding cDNA.  
 XX  
 DE Psoriasis; immunological response; anti-idiotypic antibody; glutamate;  
 KW chronic plaque; pustular; plaque-type psoriasis; psoriatic arthritis;  
 KW monoclonal antibody; 1A7; ss.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9925380-A2.  
 XX  
 PD 27-MAY-1999.  
 XX  
 PF 17-NOV-1998; 98WO-0524607.  
 XX  
 PR 16-NOV-1998; 98US-0192838.  
 PR 17-NOV-1997; 97US-0065774.  
 XX

PA (KENT ) UNIV KENTUCKY RES FOUND.  
 XX  
 PI Chatterjee M, Foon KA;  
 XX  
 DR WPI; 1999-347407/29.  
 DR P-PSDB; AAY21546.  
 XX  
 PT Treatment of psoriasis  
 XX  
 PS Disclosure; Fig 3; 48pp; English.  
 CC The invention provides a method of treating of psoriasis by administering  
 CC an antigen which has similar immunogenic properties to an antigen  
 CC expressed on cells of psoriatic tissue so that an immunological response  
 CC is elicited in the individual. The antigen stimulates the generation of  
 CC anti-idiotypic antibodies that neutralize the aberrant immune response  
 CC causing the psoriasis. The method is used to treat psoriasis, especially  
 CC chronic plaque, glutamate, pustular, plaque-type psoriasis or psoriatic  
 CC arthritis. The compositions allow the individual's own immune system to  
 CC act against psoriatic tissue. The present sequence represents a cDNA  
 CC encoding the heavy chain variable region of monoclonal antibody 1A7.  
 XX  
 SQ Sequence 458 BP; 106 A; 132 C; 113 G; 107 T; 0 other;  
 Query Match 16.5%; Score 76; DB 20; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-25;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 366 TGCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCCAAACGACACC 425  
 DB 363 TGCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCCAAACGACACC 422  
 QY 426 CCCACCCGCTCTATCCA 441  
 DB 423 CCCACCCGCTCTATCCA 438  
 RESULT 8  
 AAQ81500  
 ID AAQ81500 standard; cDNA; 861 BP.  
 XX  
 AC AAQ81500;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 30-AUG-1995 (first entry)  
 XX  
 DE SFV anti-rev sequence.  
 XX  
 KW HIV-1; human immunodeficiency virus type 1; AIDS; Rev protein;  
 KW intracellular immunization; gene therapy; single chain antibody; FV;  
 KW SFV; antibody engineering; resistance; cell immunity; HeLa; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9503832-A1.  
 XX  
 PD 09-FEB-1995.  
 XX  
 PF 28-JUL-1994; 94WO-US08448.  
 XX  
 PR 30-JUL-1993; 93US-0099870.  
 XX  
 PA (UYJE-) UNIV JEFFERSON THOMAS.  
 XX  
 PI Duan L, Pomerantz R;  
 XX  
 DR WPI; 1995-082039/11.  
 XX  
 PT Method for conducting gene therapy - comprises using recombinant  
 PT gene encoding antibody binding antigen associated with a disease;  
 PT useful for providing cell immunity.  
 XX  
 PS Example 4; Page 23; 62pp; English.



CC is useful for expressing of streptavidin fusion proteins. In particular,  
CC these are useful as tools for medical diagnostics and therapeutic  
CC purposes, e.g. for detecting the presence or absence of, or treating, a  
CC target site within a mammalian host.

XX Sequence 765 BP; 169 A; 201 C; 230 G; 165 T; 0 other;  
SQ

Query Match 13.0%; Score 60; DB 22; Length 765;  
Best Local Similarity 100.0%; Pred. No. 7.5e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ACCAGTTACAATATGCACTGGTAAAGCAGACACCTGGACAGCCCTGGAATGGATTGGA 204  
|||||  
Db 88 ACCAGTTACAATATGCACTGGTAAAGCAGACACCTGGACAGCCCTGGAATGGATTGGA 147

RESULT 11  
AAC86591  
ID AAC86591 standard; DNA; 765 BP.

XX  
AC  
XX  
AC  
XX  
DT  
XX  
DE  
XX  
DE  
XX  
DE  
XX  
KW  
XX  
KW  
XX  
OS  
XX  
OS  
XX  
OS  
XX  
PN  
XX  
PN  
XX  
PD  
XX  
PD  
XX  
PF  
XX  
PF  
XX  
PR  
XX  
PR  
XX  
PA  
XX  
PA  
XX  
PI  
XX  
PI  
XX  
DR  
XX  
DR  
XX  
PT  
XX  
PT  
XX  
PT  
XX  
PS  
XX  
PS  
XX

AAC86591;  
02-APR-2001 (first entry)  
DNA encoding a fusion of a single chain antibody and streptavidin.  
Streptavidin; tumour cell; cancer; adenocarcinoma;  
hematological malignancy; ss.  
Synthetic.  
Streptomyces avidinii.  
Unidentified.  
WO200075333-A1.  
14-DEC-2000.  
05-JUN-2000; 2000WO-US15595.  
07-JUN-1999; 99US-0137900.  
03-DEC-1999; 99US-0168976.  
(NEOR-) NEORX CORP.  
Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;  
WPI; 2001-091213/10.  
New vector constructs for expressing genomic streptavidin fusion  
proteins which are useful for targeting tumour cells associated with  
cancer, e.g. adenocarcinomas -  
Example 5; Page 95; 100pp; English.  
The present sequence encodes a fusion of an anti-CD20 single chain  
antibody and streptavidin. The fusion protein is expressed using  
vectors of the invention. The specification describes vector constructs  
for expressing streptavidin fusion proteins. The vector comprises a  
nucleic acid encoding streptavidin or its functional variant operatively  
linked to a promoter, and a cloning site for insertion of a second  
nucleic acid sequence encoding a polypeptide to be fused with  
streptavidin. Alternatively, the vector construct comprises a nucleic  
acid sequence, interposed between the promoter and the first nucleic  
acid, operatively linked to a promoter, encoding a polypeptide to be  
fused with streptavidin, and a cloning site for insertion of a second  
nucleic acid encoding at least 129 amino acids of streptavidin or its  
functional variant. The fusion proteins are useful for targeting tumour  
cells, particularly tumour cells associated with cancer,  
e.g. adenocarcinomas or hematological malignancies. The vector construct  
is useful for expressing of streptavidin fusion proteins. In particular,  
these are useful as tools for medical diagnostics and therapeutic  
purposes, e.g. for detecting the presence or absence of, or treating, a  
target site within a mammalian host.

XX  
SQ Sequence 765 BP; 170 A; 201 C; 231 G; 163 T; 0 other;  
Query Match 13.0%; Score 60; DB 22; Length 765;  
Best Local Similarity 100.0%; Pred. No. 7.5e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ACCAGTTACAATATGCACTGGTAAAGCAGACACCTGGACAGCCCTGGAATGGATTGGA 204  
|||||  
Db 88 ACCAGTTACAATATGCACTGGTAAAGCAGACACCTGGACAGCCCTGGAATGGATTGGA 147

RESULT 12  
AAC86563  
ID AAC86563 standard; DNA; 1239 BP.

XX  
AC  
XX  
AC  
XX  
DT  
XX  
DE  
XX  
DE  
XX  
DE  
XX  
KW  
XX  
KW  
XX  
OS  
XX  
OS  
XX  
OS  
XX  
PN  
XX  
PN  
XX  
PD  
XX  
PD  
XX  
PF  
XX  
PF  
XX  
PR  
XX  
PR  
XX  
PA  
XX  
PA  
XX  
PI  
XX  
PI  
XX  
DR  
XX  
DR  
XX  
PT  
XX  
PT  
XX  
PT  
XX  
PS  
XX  
PS  
XX

AAC86563;  
02-APR-2001 (first entry)  
DNA encoding a fusion of anti-CD20 single chain antibody/streptavidin.  
Streptavidin; tumour cell; cancer; adenocarcinoma;  
hematological malignancy; ss.  
Synthetic.  
Streptomyces avidinii.  
Homo sapiens.  
WO200075333-A1.  
14-DEC-2000.  
05-JUN-2000; 2000WO-US15595.  
07-JUN-1999; 99US-0137900.  
03-DEC-1999; 99US-0168976.  
(NEOR-) NEORX CORP.  
Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;  
WPI; 2001-091213/10.  
P-PSDB; AAB30694.  
New vector constructs for expressing genomic streptavidin fusion  
proteins which are useful for targeting tumour cells associated with  
cancer, e.g. adenocarcinomas -  
Example 2; Fig 11A; 100pp; English.  
The present sequence encodes a fusion of an anti-CD20 single chain  
antibody (B9E9) streptavidin. The fusion protein is expressed using  
vectors of the invention. The specification describes vector constructs  
for expressing streptavidin fusion proteins. The vector comprises a  
nucleic acid encoding streptavidin or its functional variant operatively  
linked to a promoter, and a cloning site for insertion of a second  
nucleic acid sequence encoding a polypeptide to be fused with  
streptavidin, interposed between the promoter and the first nucleic  
acid sequence. Alternatively, the vector construct comprises a nucleic  
acid, operatively linked to a promoter, encoding a polypeptide to be  
fused with streptavidin, and a cloning site for insertion of a second  
nucleic acid encoding at least 129 amino acids of streptavidin or its  
functional variant. The fusion proteins are useful for targeting tumour  
cells, particularly tumour cells associated with cancer,  
e.g. adenocarcinomas or hematological malignancies. The vector construct  
is useful for expressing of streptavidin fusion proteins. In particular,  
these are useful as tools for medical diagnostics and therapeutic  
purposes, e.g. for detecting the presence or absence of, or treating, a  
target site within a mammalian host.

XX  
SQ Sequence 1239 BP; 270 A; 392 C; 356 G; 221 T; 0 other;

Query Match 13.0%; Score 60; DB 22; Length 1239;  
Best Local Similarity 100.0%; Pred. No. 7.2e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ACCAGTTACAAATATGCACTGGGTAAGACACACACCTGGACAGGCGCTGGGAATGGATTGGA 204  
|||||  
Db 469 ACCAGTTACAAATATGCACTGGGTAAGACACACACCTGGACAGGCGCTGGGAATGGATTGGA 528  
|||||

RESULT 13  
AAC86564  
ID AAC86564 standard; DNA; 1280 BP.  
XX AC AAC86564;  
XX DT 02-APR-2001 (first entry)  
XX DE Anti-CD20 single chain antibody/streptavidin fusion protein cassette.  
XX KW Streptavidin; tumour cell; cancer; adenocarcinoma;  
XX KW hematological malignancy; ss.  
XX OS Synthetic.  
XX OS Streptomyces avidinii.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX CDS 3..1274  
FT FT /\*tag= a  
FT FT /product= "anti-CD20 scFv and streptavidin fusion"  
XX PN WO200075333-A1.  
XX XX  
XX PD 14-DEC-2000.  
XX XX  
XX PF 05-JUN-2000; 2000WO-US15595.  
XX XX  
XX PR 07-JUN-1999; 99US-0137900.  
XX PR 03-DEC-1999; 99US-0168976.  
XX XX  
XX PA (NEOR-) NEORX CORP.  
XX XX  
XX PI Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;  
XX XX  
XX DR WPI; 2001-091213/10.  
XX DR P-PSDB; AAB30695.  
XX XX  
XX PT New vector constructs for expressing genomic streptavidin fusion  
XX PT proteins which are useful for targeting tumour cells associated with  
XX PT cancer, e.g. adenocarcinomas -  
XX XX  
XX PS Example 2; Fig 11C; 100pp; English.  
XX XX  
XX CC The present sequence encodes a fusion of an anti-CD20 single chain  
XX CC antibody (B9E9) streptavidin. The fusion protein is expressed using  
XX CC vectors of the invention. The specification describes vector constructs  
XX CC for expressing streptavidin fusion proteins. The vector comprises a  
XX CC nucleic acid encoding streptavidin or its functional variant operatively  
XX CC linked to a promoter, and a cloning site for insertion of a second  
XX CC nucleic acid sequence encoding a polypeptide to be fused with  
XX CC streptavidin, interposed between the promoter and the first nucleic  
XX CC acid sequence. Alternatively, the vector construct comprises a nucleic  
XX CC acid, operatively linked to a promoter, encoding a polypeptide to be  
XX CC fused with streptavidin, and a cloning site for insertion of a second  
XX CC nucleic acid encoding at least 129 amino acids of streptavidin or its  
XX CC functional variant. The fusion proteins are useful for targeting tumour  
XX CC cells, particularly tumour cells associated with cancer,  
XX CC e.g. adenocarcinomas or hematological malignancies. The vector construct  
XX CC is useful for expressing of streptavidin fusion proteins. In particular,  
XX CC these are useful as tools for medical diagnostics and therapeutic  
XX CC purposes, e.g. for detecting the presence or absence of, or treating, a  
XX CC target site within a mammalian host.

SQ Sequence 1280 BP; 267 A; 397 C; 388 G; 228 T; 0 other;  
Query Match 13.0%; Score 60; DB 22; Length 1280;  
Best Local Similarity 100.0%; Pred. No. 7.2e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ACCAGTTACAAATATGCACTGGGTAAGACACACACCTGGACAGGCGCTGGGAATGGATTGGA 204  
|||||  
Db 96 ACCAGTTACAAATATGCACTGGGTAAGACACACACCTGGACAGGCGCTGGGAATGGATTGGA 155  
|||||

RESULT 14  
AA15019  
ID AA15019 standard; DNA; 1925 BP.  
XX AC AA15019;  
XX DT 21-AUG-2000 (first entry)  
XX DE DNA encoding a CD-20 specific chimeric receptor.  
XX KW CD20-specific receptor; CD-20 specific redirected T cell; leukemia;  
XX KW CD20+ malignancy; non-Hodgkin's lymphoma; myeloablative chemotherapy;  
XX KW stem cell rescue; autoimmune disease; lupus; rheumatoid arthritis; ss.  
XX OS Synthetic.  
XX OS Mus sp.  
XX FH Key Location/Qualifiers  
XX CDS 27..1928  
FT FT /\*tag= a  
XX PN WO200023573-A2.  
XX XX  
XX PD 27-APR-2000.  
XX XX  
XX PF 20-OCT-1999; 99WO-US24484.  
XX XX  
XX PR 20-OCT-1998; 98US-0105014.  
XX XX  
XX PA (CITY ) CITY OF HOPE.  
XX XX  
XX PI Raubitschek A, Jensen MC, Wu AM;  
XX XX  
XX DR WPI; 2000-339676/29.  
XX DR P-PSDB; AAY84965.  
XX XX  
XX PT Genetically engineered CD20-specific redirected T cells useful for  
XX PT treating a CD20+ malignancy, such as non-Hodgkin's lymphoma or CD20+  
XX PT acute or chronic leukemia, and autoimmune disease -  
XX XX  
XX PS Example 1; Page 50-53; 58pp; English.  
XX XX  
XX CC The present sequence encodes a synthetic CD20-specific chimeric  
XX CC receptor. The specification describes CD-20 specific redirected T cells  
XX CC which express and bear on the cell surface membrane a CD20-chimeric  
XX CC receptor comprising an intracellular signalling domain, a transmembrane  
XX CC domain and an extracellular domain, the extracellular domain comprising  
XX CC a CD20-specific receptor. The genetically engineered CD20-specific  
XX CC redirected T cells are useful for treating a CD20+ malignancy, such  
XX CC as non-Hodgkin's lymphoma or CD20+ acute or chronic leukemia, in a  
XX CC human patient having previously undergone myeloablative chemotherapy and  
XX CC stem cell rescue. The genetically engineered CD20-specific redirected  
XX CC T cells are also useful for abrogating an untoward B cell function, such  
XX CC as autoimmune disease (lupus or rheumatoid arthritis) in a patient.  
XX XX  
XX SQ Sequence 1925 BP; 471 A; 554 C; 541 G; 359 T; 0 other;  
Query Match 13.0%; Score 60; DB 21; Length 1925;  
Best Local Similarity 100.0%; Pred. No. 7.1e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ACCAGTTACAAATATGCACTGGGTAAGACACACACCTGGACAGGCGCTGGGAATGGATTGGA 204



Job time : 203.688 secs

Db 546 ACCAGTTACATATGCTAGCTGGTAAAGACAGACACCTTGGACAGGGCCTGGAATGATTGGA 605  
|||||

RESULT 15

AAT96345

ID AAT96345 standard; cDNA; 360 BP.

XX AC AAT96345;

XX XX

DT 08-APR-1998 (first entry)

XX XX

DE cDNA for Ig heavy chain variable region of anti-asparaginase MAb.

XX XX

KW Immunoglobulin; Ig; heavy chain; variable region; murine;

KW asparaginase II; monoclonal antibody; MAb;

KW recombinant chimeric polypeptide; ss.

XX XX

OS Mus sp.

XX XX

PN US5686579-A.

XX XX

PD 11-NOV-1997.

XX XX

PF 23-MAY-1995; 95US-0447422.

XX XX

PR 22-JUN-1993; 93US-0081410.

PR 21-JUN-1988; 88US-0205748.

PR 31-AUG-1992; 92US-0938505.

PR 23-MAY-1995; 95US-0447422.

XX XX

PA (HYBR-) HYBRISSENS LTD.

XX XX

PI Ramjeesingh M, Rothstein A, Shami EY;

XX XX

DR WPI; 1997-558200/51.

XX XX

PT Self-protecting chimeric polypeptide comprising biologically active

PT sequence and single-chain antibody sequence - has resistance to e.g.

PT disrupting temperature, presence of proteolytic enzymes, etc.

XX XX

PS Example 2; Columns 27-28; 29pp; English.

XX XX

CC The present sequence is the cDNA for the immunoglobulin (Ig) heavy

CC chain variable region of a murine anti-asparaginase II monoclonal

CC antibody (MAB). The cDNA was used in the preparation of a novel

CC recombinant chimeric polypeptide, comprising a 1st region

CC comprising a biologically active domain and another domain

CC containing an epitope, linked via a polypeptide to a 2nd region

CC including a single chain antibody (SCA) having the light and heavy

CC chains of an antibody variable region which specifically binds the

CC epitope in the 1st region. The chimeric polypeptide assumes a

CC conformation in which the SCA is bound to the epitope of the 1st

CC region and protects its biological activity from deactivation by

CC denaturing temperatures or pH conditions, proteolytic enzymes,

CC oxidising agents or alcohol. The regions of the chimeric

CC polypeptide interact to form a structure analogous to an

CC antibody-antigen complex. A L-asparaginase-SCA fusion protein of

CC the above type has better trypsin resistance than free

CC L-asparaginase.

XX XX

SQ Sequence 360 BP; 85 A; 101 C; 95 G; 79 T; 0 other;

Query Match

Best Local Similarity 12.8%; Score 59; DB 18; Length 360;

Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 TGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCCAGCCAAACGACACCCCA 429

DB 270 TGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCCAGCCAAACGACACCCCA 328

Search completed: August 30, 2003, 19:57:04

**THIS PAGE BLANK (uspto)**

Result No.	Query			ID	Description
	Score	Match	Length		
1	461	100.0	461	US-09-861-294-3	Sequence 3, Appli
2	461	100.0	461	US-10-367-506-3	Sequence 3, Appli
3	76	16.5	458	US-09-990-205-3	Sequence 3, Appli
4	76	16.5	458	US-10-153-401-3	Sequence 3, Appli
5	60	13.0	765	US-10-244-821-44	Sequence 44, Appli
6	60	13.0	765	US-10-244-821-44	Sequence 44, Appli
7	60	13.0	765	US-10-013-173-44	Sequence 44, Appli
8	60	13.0	765	US-10-013-173-44	Sequence 44, Appli
9	60	13.0	765	US-10-150-762-44	Sequence 44, Appli
10	60	13.0	765	US-10-150-762-44	Sequence 44, Appli
11	60	13.0	1239	US-10-244-821-5	Sequence 5, Appli
12	60	13.0	1239	US-10-013-173-5	Sequence 5, Appli
13	60	13.0	1239	US-10-150-762-5	Sequence 5, Appli
14	60	13.0	1280	US-10/244	Sequence 7, Appli
15	60	13.0	1280	US-10/1013	Sequence 7, Appli
16	60	13.0	1280	US-10/150	Sequence 7, Appli



GENERAL INFORMATION:  
APPLICANT: Chatterjee, Malaya  
Foon, Kenneth A.  
Chatterjee, Sunil K.  
TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE  
TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/153,401  
FILING DATE: 27-Aug-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/293,533  
FILING DATE: 1999-04-15  
APPLICATION NUMBER: US 08/372,676  
FILING DATE: 1995-01-17  
APPLICATION NUMBER: US 08/591,196  
FILING DATE: 1996-01-16  
ATTORNEY/AGENT INFORMATION:  
NAME: Catherine M. Polizzi  
REGISTRATION NUMBER: 40,130  
REFERENCE/DOCKET NUMBER: 304142000202  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 458 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..456  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 58  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-153-401-3  
Query Match 16.5%; Score 76; DB 14; Length 458;  
Best Local Similarity 100.0%; Pred. No. 2.6e-30;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 366 TGCTCTGGACTACTGGGGTCAAGGACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 425  
Db 363 TGCTCTGGACTACTGGGGTCAAGGACCTCAGTCACCGTCTCCTCAGCCAAACGACACC 422  
QY 426 CCCACCCGCTATCCA 441  
Db 423 CCCACCCGCTATCCA 438  
RESULT 5  
US-10-244-821-44  
; Sequence 44, Application US/10244821  
; Publication No. US20030143233A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen Charles

APPLICANT: Graves, Scott Stoll  
APPLICANT: Schultz, Joanne Elaine  
APPLICANT: Lin, Yukang  
APPLICANT: Sanderson, James Allen  
APPLICANT: Reno, John M.  
APPLICANT: Dearstyne, Erica A.  
TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
METHODS OF USE THEREOF  
FILE REFERENCE: 690022.547C3  
CURRENT APPLICATION NUMBER: US/10/244,821  
CURRENT FILING DATE: 2002-09-16  
NUMBER OF SEQ ID NOS: 92  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 44  
LENGTH: 765  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: B9E9 single chain antibody-genomic Streptavidin  
OTHER INFORMATION: fusion construct  
US-10-244-821-44  
Query Match 13.0%; Score 60; DB 12; Length 765;  
Best Local Similarity 100.0%; Pred. No. 8.8e-22;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 145 ACCAGTTACAATATGCACCTGGGTAAAGCAGACACCTGGAGCGCCTGGATGGGA 204  
Db 88 ACCAGTTACAATATGCACCTGGGTAAAGCAGACACCTGGAGCGCCTGGATGGGA 147  
RESULT 6  
US-10-244-821-45  
; Sequence 45, Application US/10244821  
; Publication No. US20030143233A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen Charles  
; APPLICANT: Graves, Scott Stoll  
; APPLICANT: Schultz, Joanne Elaine  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James Allen  
; APPLICANT: Reno, John M.  
; APPLICANT: Dearstyne, Erica A.  
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
; METHODS OF USE THEREOF  
; FILE REFERENCE: 690022.547C3  
; CURRENT APPLICATION NUMBER: US/10/244,821  
; CURRENT FILING DATE: 2002-09-16  
; NUMBER OF SEQ ID NOS: 92  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 45  
; LENGTH: 765  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: C2B8 single chain antibody-genomic Streptavidin  
; OTHER INFORMATION: fusion construct  
; US-10-244-821-45  
Query Match 13.0%; Score 60; DB 12; Length 765;  
Best Local Similarity 100.0%; Pred. No. 8.8e-22;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 145 ACCAGTTACAATATGCACCTGGGTAAAGCAGACACCTGGAGCGCCTGGATGGGA 204  
Db 88 ACCAGTTACAATATGCACCTGGGTAAAGCAGACACCTGGAGCGCCTGGATGGGA 147  
RESULT 7  
US-10-013-173-44  
; Sequence 44, Application US/10013173  
; Publication No. US20030095977A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen Charles

; APPLICANT: Goshorn, Stephen C.  
; APPLICANT: Graves, Scott Stoll  
; APPLICANT: Schultz, Joanne Elaine  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James A.  
; APPLICANT: Reno, Jonh M.  
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 690022.547C1  
; CURRENT APPLICATION NUMBER: US/10/013.173  
; CURRENT FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 765  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: B9E9 single chain antibody-genomic Streptavidin  
; OTHER INFORMATION: fusion construct  
US-10-013-173-44

Query Match 13.0%; Score 60; DB 14; Length 765;  
Best Local Similarity 100.0%; Pred. No. 8.8e-22;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ACCAGTTACAAATATGCACCTGGGTAAGACACACACCTGGAGCGCCTGGGAATGGATTGGA 204  
|||||  
DB 88 ACCAGTTACAAATATGCACCTGGGTAAGACACACACCTGGAGCGCCTGGGAATGGATTGGA 147  
|||||

RESULT 8  
US-10-013-173-45  
; Sequence 45, Application US/10013173  
; Publication No. US2003009577A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen C.  
; APPLICANT: Graves, Scott Stoll  
; APPLICANT: Schultz, Joanne Elaine  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James A.  
; APPLICANT: Reno, Jonh M.  
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 690022.547C1  
; CURRENT APPLICATION NUMBER: US/10/013.173  
; CURRENT FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45  
; LENGTH: 765  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: C2B8 single chain antibody-genomic Streptavidin  
; OTHER INFORMATION: fusion construct  
US-10-013-173-45

Query Match 13.0%; Score 60; DB 14; Length 765;  
Best Local Similarity 100.0%; Pred. No. 8.8e-22;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ACCAGTTACAAATATGCACCTGGGTAAGACACACACCTGGAGCGCCTGGGAATGGATTGGA 204  
|||||  
DB 88 ACCAGTTACAAATATGCACCTGGGTAAGACACACACCTGGAGCGCCTGGGAATGGATTGGA 147  
|||||

RESULT 9  
US-10-150-762-44  
; Sequence 44, Application US/10150762  
; Publication No. US20030103948A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen C.

; APPLICANT: Graves, Scott S.  
; APPLICANT: Schultz, Joanne E.  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James A.  
; APPLICANT: Reno, Jonh M.  
; APPLICANT: Dearstynne, Erica A.  
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 690022.547C2  
; CURRENT APPLICATION NUMBER: US/10/150.762  
; CURRENT FILING DATE: 2002-05-17  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 765  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: B9E9 single chain antibody-genomic Streptavidin  
; OTHER INFORMATION: fusion construct  
US-10-150-762-44

Query Match 13.0%; Score 60; DB 14; Length 765;  
Best Local Similarity 100.0%; Pred. No. 8.8e-22;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ACCAGTTACAAATATGCACCTGGGTAAGACACACACCTGGAGCGCCTGGGAATGGATTGGA 204  
|||||  
DB 88 ACCAGTTACAAATATGCACCTGGGTAAGACACACACCTGGAGCGCCTGGGAATGGATTGGA 147  
|||||

RESULT 10  
US-10-150-762-45  
; Sequence 45, Application US/10150762  
; Publication No. US20030103948A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen C.  
; APPLICANT: Graves, Scott S.  
; APPLICANT: Schultz, Joanne E.  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James A.  
; APPLICANT: Reno, Jonh M.  
; APPLICANT: Dearstynne, Erica A.  
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 690022.547C2  
; CURRENT APPLICATION NUMBER: US/10/150.762  
; CURRENT FILING DATE: 2002-05-17  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45  
; LENGTH: 765  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: C2B8 single chain antibody-genomic Streptavidin  
; OTHER INFORMATION: fusion construct  
US-10-150-762-45

Query Match 13.0%; Score 60; DB 14; Length 765;  
Best Local Similarity 100.0%; Pred. No. 8.8e-22;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ACCAGTTACAAATATGCACCTGGGTAAGACACACACCTGGAGCGCCTGGGAATGGATTGGA 204  
|||||  
DB 88 ACCAGTTACAAATATGCACCTGGGTAAGACACACACCTGGAGCGCCTGGGAATGGATTGGA 147  
|||||

RESULT 11  
US-10-244-821-5  
; Sequence 5, Application US/10244821  
; Publication No. US20030143233A1  
; GENERAL INFORMATION:

RESULT 13  
US-10-150-762-5  
; Sequence 5, Application US/10150762  
; Publication No. US20030103948A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen C..  
; APPLICANT: Graves, Scott S.

```

RESULT 15
US/100/013
; Sequence 7, Application US/10013173
; Publication No. US2003009577A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine

```

```
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, Jonh M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1280
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic streptavidin fusion construct
US/10/013,173-7

Query Match      13.0%; Score 60; DB 14; Length 1280;
Best Local Similarity 100.0%; Pred. No. 8.7e-22;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      145 ACCAGTTACAATATGCACCTGGGTAAAGCAGACACACCTGGAGGGCCTGGGAATGGATTGGA 204
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       96 ACCAGTTACAATATGCACCTGGGTAAAGCAGACACACCTGGAGGGCCTGGGAATGGATTGGA 155

Search completed: August 30, 2003, 22:02:42
Job time : 168.244 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 18:52:40 ; Search time 1604.24 Seconds  
(without alignments)  
6984.224 Million cell updates/sec

Title: US-08-836-455-3  
Perfect score: 461  
Sequence: 1 ATGGAATCAGCTGGGTCTT.....CTGGTCCCTGGAAGCTTGGG 461

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estinu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_tod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	12.8	676	10	BE369087
2	49	10.6	639	10	BE371136
3	49	10.6	802	10	BE137216
4	49	10.6	861	12	BI455668

5	49	10.6	883	12	BI151077
6	49	10.6	1012	10	BF142302
7	48	10.4	718	10	BF136279
8	48	10.4	862	10	BF143948
9	47	10.2	913	10	BF583109
10	47	10.2	933	13	BQ929209
11	47	10.2	991	14	BY703373
12	47	10.2	1384	11	AK002875
13	43	9.3	406	10	BF016722
14	42	9.1	429	14	CA579143
15	42	9.1	451	14	CA580198
16	42	9.1	458	14	CA580083
17	42	9.1	465	14	CA578116
18	42	9.1	473	14	CA580167
19	42	9.1	475	14	CA578534
20	42	9.1	477	14	CA570382
21	42	9.1	480	14	CA577562
22	42	9.1	485	14	CA577844
23	42	9.1	496	14	CA578504
24	42	9.1	506	14	CA578474
25	42	9.1	506	14	CA578686
26	42	9.1	519	14	CA580087
27	42	9.1	531	14	CA571287
28	42	9.1	540	14	CA579750
29	42	9.1	560	10	BF581663
30	42	9.1	567	13	BQ266826
31	42	9.1	604	10	BE367979
32	42	9.1	650	10	BF579001
33	42	9.1	666	12	BG963642
34	42	9.1	671	10	BF182141
35	42	9.1	690	12	BG968682
36	42	9.1	706	10	BE915586
37	42	9.1	724	10	BF168514
38	42	9.1	738	12	BG965088
39	42	9.1	746	12	BG966840
40	42	9.1	776	10	BF182218
41	42	9.1	781	12	BG967007
42	42	9.1	811	12	BG966605
43	42	9.1	819	12	BG966397
44	42	9.1	837	12	BG966355
45	42	9.1	839	13	BQ952276

ALIGNMENTS

RESULT 1  
BE369087

LOCUS  
DEFINITION

ACCESSION  
VERSION

KEYWORDS  
SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE369087 601221765F1 NCI\_CGAP\_Lu29 Mus musculus cdna clone IMAGE:3590320 5',  
676 bp mRNA linear EST 21-JUL-2000  
mrna sequence.  
BE369087  
BE369087.1 GI:9314450  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 676)  
NIH-MGC <http://mgi.nci.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapps-r@mail.nih.gov](mailto:cgapps-r@mail.nih.gov)  
Tissue procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLM8758 row: b column: 17  
High quality sequence stop: 580.

```

FEATURES
source      Location/Qualifiers
1. .676
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:3590320"
/tissue_type="spontaneous tumor, metastatic to mammary.
stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: Salt;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      161 a 188 c 173 g 154 t
ORIGIN

Query Match      12.8%; Score 59; DB 10; Length 676;
Best Local Similarity 100.0%; Pred. No. 1.5e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 TGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCAAAACGACACCCCA 429
|||||
390 TGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCAAAACGACACCCCA 448

RESULT 2
BE3711136      639 bp mRNA linear EST 21-JUL-2000
LOCUS
DEFINITION
601218628F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3587742 5',
mRNA sequence.
ACCESSION
BE3711136
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 639)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8751 row: q column: 07
High quality sequence stop: 552.
Location/Qualifiers
1. .639
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:3587742"
/tissue_type="spontaneous tumor, metastatic to mammary.
stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: Salt;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      149 a 171 c 164 g 155 t
ORIGIN

Query Match      10.6%; Score 49; DB 10; Length 639;
Best Local Similarity 100.0%; Pred. No. 2.8e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 TGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCAAAAC 419
|||||
433 TGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCAAAAC 481

RESULT 4
BI455668      861 bp mRNA linear EST 21-AUG-2001
LOCUS
DEFINITION
603173862F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5253279 5',
mRNA sequence.
ACCESSION
BI455668
VERSION
BI455668.1 GI:15246324
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse).
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Best Local Similarity 100.0%; Pred. No. 2.6e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 TGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCAAAAC 419
|||||
399 TGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCAAAAC 447

RESULT 3
BF137216      802 bp mRNA linear EST 24-OCT-2000
LOCUS
DEFINITION
601784466F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4012360 5',
mRNA sequence.
ACCESSION
BF137216
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 802)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9253 row: c column: 17
High quality sequence stop: 697.
Location/Qualifiers
1. .802
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4012360"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; transgenic model WNT-1, expression driven by
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dt. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT      191 a 217 c 211 g 183 t
ORIGIN

Query Match      10.6%; Score 49; DB 10; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.8e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 TGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCAAAAC 419
|||||
433 TGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCAAAAC 481

RESULT 4
BI455668      861 bp mRNA linear EST 21-AUG-2001
LOCUS
DEFINITION
603173862F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5253279 5',
mRNA sequence.
ACCESSION
BI455668
VERSION
BI455668.1 GI:15246324
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse).
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

round through the I.M.R.G.E. Consortium/LEML at:
plate://image.llnl.gov
Plate: LRAM9279 row: h column: 12
High quality sequence stop: 711.
Location/Qualifiers
1. .1012
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4022459"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; transgenic model WNT-1, expression driven by
MTWV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dt. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 273 a 271 c 267 g 201 t
ORIGIN
Query Match 10.6%; Score 49; DB 10; Length 1012;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 371 TGGACTACTGGGGTCAAGAACCTCAGTCACCGTCTCTCTCAGCCAAAC 419  
 LOCUS  
 Db 411 TGGACTACTGGGGTCAAGAACCTCAGTCACCGTCTCTCTCAGCCAAAC 459

RESULT 7  
 BF136279 601780989F1 NCI\_CGAP\_Lu30 718 bp mRNA linear EST 24-OCT-2000  
 DEFINITION  
 mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

BF136279.1 GI:10975319  
 EST.  
 Mus musculus (house mouse)

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 718)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM9244 row: 0 column: 22

High quality sequence stop: 671.

Location/Qualifiers

source

1..718

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="IMAGE:4009197"

/tissue\_type="tumor, metastatic to mammary"

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP\_Lu30"

/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: NotI;

Site\_2: SalI; transgenic model WNT-1, expression driven by

MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo

dt. Library constructed by Life Technologies.

Investigator providing samples: Gilbert Smith, NIH"

173 a 196 c 181 g 168 t

BASE COUNT

ORIGIN

Query Match 10.4%; Score 48; DB 10; Length 718;

Best Local Similarity 100.0%; Pred. No. 8.9e-14;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 TGGACTACTGGGGTCAAGAACCTCAGTCACCGTCTCTCTCAGCCAAAC 418

Db 410 TGGACTACTGGGGTCAAGAACCTCAGTCACCGTCTCTCTCAGCCAAAC 457

RESULT 8

LOCUS

BF143948 601780493F1 NCI\_CGAP\_Lu30 862 bp mRNA linear EST 24-OCT-2000

DEFINITION

mRNA sequence.

ACCESSION

BF143948

BF143948.1 GI:10982988

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 862)

AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM9258 row: 1 column: 23

High quality sequence stop: 671.

Location/Qualifiers

source

1..862

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="IMAGE:4014430"

/tissue\_type="tumor, metastatic to mammary"

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP\_Lu30"

/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: NotI;

Site\_2: SalI; transgenic model WNT-1, expression driven by

MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo

dt. Library constructed by Life Technologies.

Investigator providing samples: Gilbert Smith, NIH"

213 a 236 c 236 g 176 t

BASE COUNT

ORIGIN

Query Match 10.4%; Score 48; DB 10; Length 862;

Best Local Similarity 100.0%; Pred. No. 9.5e-14;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 372 GGACTACTGGGGTCAAGAACCTCAGTCACCGTCTCTCTCAGCCAAAC 419

Db 411 GGACTACTGGGGTCAAGAACCTCAGTCACCGTCTCTCTCAGCCAAAC 458

RESULT 9

LOCUS

BF583109 602098016F1 NCI\_CGAP\_Co24 913 bp mRNA linear EST 12-DEC-2000

DEFINITION

mRNA sequence.

ACCESSION

BF583109

BF583109.1 GI:11656827

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 913)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM9797 row: p column: 04

High quality sequence stop: 656.

Location/Qualifiers

source

1..913

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

```

/db_xref="taxon:10090"
/clone="IMAGE:4218099"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      183 a      261 c      250 g      219 t
ORIGIN
Query Match      10.2%; Score 47; DB 10; Length 913;
Best Local Similarity 100.0%; Pred. No. 3.2e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 366 TGCTCGGACTACTGGGGTCAAGAACCTCAGTCACCGTCTCCTCAG 412
|||||
Db 383 TGCTCGGACTACTGGGGTCAAGAACCTCAGTCACCGTCTCCTCAG 429
|||||

RESULT 10
B0929209
LOCUS      B0929209      933 bp      mRNA      linear      EST 20-AUG-2002
DEFINITION AGENCOURT_9533380 NCI_CGAP_Co24 Mus musculus cDNA clone
IMAGE:6477383 5', mRNA sequence.
B0929209
ACCESSION B0929209
VERSION   B0929209
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 933)
AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Saito,N., Hasegawa,Y., Nogami,A., Schonbach,C.,
Yagi,K., Tomaru,Y., Hasegawa,Y., Hill,D.P., Bult,C., Hume,D.A.,
Gojobori,T., Baldarelli,R., Kanapin,A., Matsuda,H., Batalov,S.,
Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani
,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest
,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A.,
Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J.,
Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M.,
King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons
,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Mikl
,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G.,
Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D.,
Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring
,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou
,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale
,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y.,
Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa
,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A.,
Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura
,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii
,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata
,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander
,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
ANALYSIS of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
22354683
PUBMED
COMMENT      Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsukuba-shi, Ibaraki, Japan 305-8565, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda
,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K.,
Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno
,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K.,
Numazaki,R., Ohno,M., Ohnishi,N., Saito,R., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y.,
Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct
Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken

```

Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

#### FEATURES

Location/Qualifiers  
 1. 991  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="0610041A01"  
 /sex="male"  
 /tissue\_type="kidney"  
 /dev\_stage="adult"  
 /lab\_host="SOLR"  
 /clone\_lib="RIKEN full-length enriched, adult male kidney"  
 /note="Site 1: XhoI; Site 2: SstI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5']  
 GAGAGAGCGCGCCCACTCGAGTCTTTTCTTTTCTTTT 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5']  
 GAGAGAGCGCGCCCACTCGAGTCTTTTCTTTTCTTTT 3'.  
 cDNA was cleaved with XhoI and SstI."  
 BASE COUNT 231 a 289 c 238 g 232 t 1 others  
 ORIGIN

Query Match 10.2%; Score 47; DB 14; Length 991;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-13;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 TCAGCAGCCTGACATCTGAAGACTCTCGGCTCTATTCTGTGCAAGA 351  
 |||||  
 DB 186 TCAGCAGCCTGACATCTGAAGACTCTCGGCTCTATTCTGTGCAAGA 232  
 |||||

RESULT 12  
 AK002875 1384 bp mRNA linear HTC 05-DEC-2002  
 LOCUS  
 DEFINITION Mus musculus adult male kidney cDNA, RIKEN full-length enriched library, clone:0610041A01 product:immunoglobulin heavy chain 6 (heavy chain of IgM), full insert sequence.

ACCESSION AK002875  
 VERSION AK002875.1 GI:12833178  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)

ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 REFERENCE 1  
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636

REFERENCE 2  
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159

REFERENCE 3  
 AUTHORS Shibata,K., Itoh,M., Alzawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,

#### TITLE

JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

Yamamoto,R., Matsumoto,H., Sakauchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Kiyosawa,H., Kondo,S., Yamana,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,Y., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Tomita,M., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Wagner,L., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 Nature 420, 563-573 (2002)  
 MEDLINE 21085660  
 PUBMED 11217851

#### TITLE

JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 MEDLINE 21085660  
 PUBMED 11217851

#### TITLE

JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:<http://genome.gsc.riken.go.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216)  
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

#### COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'] GAGAGAGCGCGCCCACTCGAGTCTTTTCTTTTCTTTT 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0. Second strand cDNA was prepared with the primer adapter of sequence [5'] GAGAGAGCGCGCCCACTCGAGTCTTTTCTTTTCTTTT 3'. cDNA was

cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:  
XhoI. Host: SOLR.

## FEATURES

Location/Qualifiers  
1. .1384  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:0610041A01"  
/db\_xref="MGI:1907162"  
/db\_xref="taxon:10090"  
/clone="0610041A01"  
/sex="male"  
/tissue\_type="kidney"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
53. .1333  
/note="unnamed protein product: Immunoglobulin heavy chain  
6 (heavy chain of IgM) (MGI:96448, GB|AK008342,  
evidence: BLASTN, 100%, match=137)  
putative"  
/codon\_start=1  
/protein\_id="BAB22422.1"  
/db\_xref="GI:12833179"  
/db\_xref="MGI:96486"  
/translation="MGFSRIFLLSVTTGNGDYSYNQKFKGKATLVDKSSSTAYWQ  
LSSITSDSAVYFCARDSDYGPYAMDYWGQSTSVTSSSARNPTIYPLTPRALSSD  
PVTGCLIHDFPSTGMVTKGSKDITVTFPPALASGGYTMSSQLTLPVACEPE  
GTVKCSVQHDNSAVQELDKVCGPPPCPCPPCHPSLSLORPALEDLLIGSDASL  
TCTLNGLRNPEGAVTWPSTGDKAVOKKAVONSCGYSVSVLPGCAERNWSCASPK  
CTVTHPSRDLTGTIAKITVTFPPQVHLLPPSEELALNELVSLTCLVRAFNPKVL  
VRMLHNEUSPESYLVFPELKEPGEGATTYLVTSVLRVSAELWKQDQVSCWVGHEA  
LPNMFTQKTLDRLSGKPTNSVSVINSEGDGICY"  
1360. .1365  
/note="putative"  
1384  
/note="putative"

## CDS

6 (heavy chain of IgM) (MGI:96448, GB|AK008342,  
evidence: BLASTN, 100%, match=137)  
putative"

polyA\_signal  
1360. .1365  
/note="putative"  
1384  
/note="putative"

## polyA\_site

321 a 413 c 342 g 308 t

## BASE COUNT

ORIGIN

Query Match 10.2%; Score 47; DB 11; Length 1384;  
Best Local Similarity 100.0%; Pred. No. 3.8e-13;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 305 TCACAGCCTGACATCTGAGACTCTCGGCTCTATTCTGTGCAAGA 351

|||||

Db 186 TCACAGCCTGACATCTGAGACTCTCGGCTCTATTCTGTGCAAGA 232

## RESULT 13

## BF016722

LOCUS uy34h12.y1 NCI\_CGAP\_Lu30 Mus musculus cDNA clone IMAGE:3661511 5',  
DEFINITION similar to SW:HV02\_MOUSE P01746 IG HEAVY CHAIN V REGION 93G7  
PRECURSOR. : mRNA sequence.

## ACCESSION

## BF016722

## VERSION

## BF016722.1

## KEYWORDS

## EST.

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
image.llnl.gov/image/html/iresources.shtml

MGI:1422279

Seq primer: -40RP from Gibco.

## FEATURES

## source

Location/Qualifiers

1. .406

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="IMAGE:3661511"

/tissue\_type="tumor, metastatic to mammary"

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP\_Lu30"

/note="Organ: lung; Vector: pCMV-SPORT6; Site:1: NotI;  
Site:2: SalI; transgenic model WNT-1, expression driven by  
MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo  
dT. Library constructed by Life Technologies.  
Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 109 a 102 c 98 g 97 t

## ORIGIN

Query Match 9.3%; Score 43; DB 10; Length 406;  
Best Local Similarity 100.0%; Pred. No. 3.1e-11;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 98 CTGGGGCTCAGTCAAGATGCTCTGCAAGGCTTCTGGCTACAC 140

|||||

Db 133 CTGGGGCTCAGTCAAGATGCTCTGCAAGGCTTCTGGCTACAC 175

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

/clone\_lib="NTA Mouse Hematopoietic Stem Cell  
(Lin-/c-Kit-/Sca-1-) cDNA library (Long)"  
/note="vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (http://lgun-grc.nia.nih.gov/cDNA). This is  
a long-transcript enriched cDNA library (Ref. genome Res.  
11: 1553-1558 (2001)). [PMID: 11544199]]. Total RNAs were  
obtained from Drs. Dennis Taub, Dan Longo (National  
Institute on Aging, USA), Jonathan Keller (National Cancer  
Institute, USA). Double-stranded cDNAs were synthesized  
with an Oligo(dT) primer [Invitrogen]:  
5'-pGACTAGTCTAGTCGCGAGCGCGCCCTTTT-3' from  
0.9 ug of total RNA, treated with T4 DNA polymerase, and  
purified by ethanol-precipitation. The cDNAs were ligated  
to Loe-linker L1-Sal4, purified by phenol/chloroform, and  
separated from free linkers by Centricon 100. Then, the  
cDNAs were amplified by long-range high fidelity PCR using  
Ex Taq polymerase (Takara) with a primer Sal4-S. The  
products were purified by phenol/chloroform and Centricon  
100. The cDNAs were digested with SalI and NotI enzymes  
and cloned into SalI/NotI site of pSPORT1 plasmid vector.  
The DH10B E. coli host was transformed with the ligation  
mixture by the standard chemical method. The average  
insert size is about 2.1 kb. The library was constructed  
by Yulan Piao (NIA)." 104 a 112 c 108 g 105 t

Query Match 9.1%; Score 42; DB 14; Length 429;  
Best Local Similarity 100.0%; Pred. No. le-10;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 371 TGGACTACTGGGTCAGGACCTCAGTCACCGTCTCCTCAG 412  
|||||  
Db 353 TGGACTACTGGGTCAGGACCTCAGTCACCGTCTCCTCAG 394

RESULT 15  
CA580198 451 bp mRNA linear EST 19-NOV-2002  
LOCUS K0748C01-5N NTA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)  
DEFINITION CDNA Library (Long) Mus musculus cDNA clone NIA:K0748C01  
IMAGE:30078168 5', mRNA sequence.  
CA580198  
CA580198.1 GI:25128589  
EST.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus

ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 451)  
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,  
Alba, K., Taub, D., Longo, D.L., Keller, J., and Ko, M.S.H.  
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell  
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)  
Unpublished  
Other ESTs: K0748C01-3  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@lgun-grc.nia.nih.gov  
Plate: K0748 row: C column: 01  
Seq primer: M13 Reverse  
High quality sequence stop: 451  
POLYA-No. Location/Qualifiers

FEATURES  
source  
1..451  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6Ncr"  
/db\_xref="niaEST:K0748C01-5N"

/db\_xref="taxon:10090"  
/clone\_lib="NTA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)"  
/tissue\_type="Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)"  
/dev\_stage="Age approx. 10 weeks old"  
/lab\_host="DH10B"  
/clone\_lib="NTA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)"  
/note="vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (http://lgun-grc.nia.nih.gov/cDNA). This is  
a long-transcript enriched cDNA library (Ref. genome Res.  
11: 1553-1558 (2001)). [PMID: 11544199]]. Total RNAs were  
obtained from Drs. Dennis Taub, Dan Longo (National  
Institute on Aging, USA), Jonathan Keller (National Cancer  
Institute, USA). Double-stranded cDNAs were synthesized  
with an Oligo(dT) primer [Invitrogen]:  
5'-pGACTAGTCTAGTCGCGAGCGCGCCCTTTT-3' from  
0.9 ug of total RNA, treated with T4 DNA polymerase, and  
purified by ethanol-precipitation. The cDNAs were ligated  
to Loe-linker L1-Sal4, purified by phenol/chloroform, and  
separated from free linkers by Centricon 100. Then, the  
cDNAs were amplified by long-range high fidelity PCR using  
Ex Taq polymerase (Takara) with a primer Sal4-S. The  
products were purified by phenol/chloroform and Centricon  
100. The cDNAs were digested with SalI and NotI enzymes  
and cloned into SalI/NotI site of pSPORT1 plasmid vector.  
The DH10B E. coli host was transformed with the ligation  
mixture by the standard chemical method. The average  
insert size is about 2.1 kb. The library was constructed  
by Yulan Piao (NIA)." 99 a 129 c 104 g 119 t

Query Match 9.1%; Score 42; DB 14; Length 451;  
Best Local Similarity 100.0%; Pred. No. 1.1e-10;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 371 TGGACTACTGGGTCAGGACCTCAGTCACCGTCTCCTCAG 412  
|||||  
Db 363 TGGACTACTGGGTCAGGACCTCAGTCACCGTCTCCTCAG 404

Search completed: August 30, 2003, 21:57:09  
Job time : 1606.24 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 13:43:39 ; Search time 2088.91 Seconds  
(without alignments)  
9028.331 Million cell updates/sec

Title: US-08-836-455-3

Perfect score: 461

Sequence: 1 ATGGAATGCAGCTGGCTCTT.....CTGGTCCCTGGAAGCTTGGG 461

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_man:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB	ID	Description
1	461	100.0	461	6	AR164506	AR164506 Sequence
2	461	100.0	461	6	BD085738	BD085738 Methods O
3	409.4	88.8	411	10	AF124720	AF124720 Mus muscu
4	355.2	77.0	471	10	AB050080	AB050080 Mus muscu
5	353.6	76.7	457	10	AY178830	AY178830 Mus muscu
6	337.8	73.3	1542	10	BC002121	BC002121 Mus muscu
7	328.2	71.2	483	10	MMU39899	U39899 Mus musculus
8	327.2	71.0	409	6	BD092129	BD092129 Diagnosti
9	325	70.5	1544	10	MUSIGB1H1	D14625 Mus musculu
10	324.6	70.4	453	6	ARI09947	ARI09947 Sequence
11	323	70.1	525	10	MUSIGHALPA	M28834 Mus musculus
12	323	70.1	540	6	A13735	A13735 variable re
13	322	69.8	458	6	I05921	I05921 Sequence 37
14	322	69.8	458	6	I08811	I08811 Sequence 12
15	322	69.8	458	6	I09199	I09199 Sequence 38
16	322	69.8	458	10	MUSIGHXW	M17953 Mouse lg re
17	321.6	69.8	450	10	AY178829	AY178829 Mus muscu
18	320.6	69.5	1553	6	E08434	E08434 cDNA encodi
19	320.2	69.5	902	12	XX049832	U49832 Synthetic s
20	320	69.4	1683	6	E35543	E35543 Transgenic
21	318.6	69.1	473	10	AB050071	AB050071 Mus muscu
22	314.8	68.3	494	6	BD021876	BD021876 Humanized
23	313.4	68.0	451	10	MMLB4IHEV	X65773 M.musculus
24	311.6	67.6	1568	6	E33134	E33134 Humanized a
25	311.6	67.6	1570	6	A22261	A22261 M.musculus
26	311.6	67.6	1570	6	A77138	A77138 Sequence 6
27	311	67.5	489	10	MMU39900	U39900 Mus musculus
28	310.8	67.4	1526	10	MMU555822	AJ555622 Mus muscu
29	310.6	67.4	1573	10	BC018280	BC018280 Mus muscu
30	310.2	67.3	1581	6	A78881	A78881 Sequence 1
31	310.2	67.3	1581	10	MMIGHC2AA	X70423 M.musculus
32	310	67.2	1570	6	AR029102	AR029102 Sequence
33	308.6	66.9	1564	10	BC019489	BC019489 Mus muscu
34	308.2	66.9	588	6	A23165	A23165 Artificial
35	308.2	66.9	588	10	MUSIGGVZ	I02346 Mus musculu
36	308	66.8	490	10	MUSIGAC11A	M24785 Mouse anti-
37	308	66.8	1558	10	AF466769	AF466769 Mus muscu
38	307	66.6	484	10	MUSIGH4C11	M54977 Mus musculu
39	306.8	66.6	520	10	MMU416332	A7416332 Mus muscu
40	306.6	66.5	2154	10	BC018315	BC018315 Mus muscu
41	306.4	66.5	1500	10	MMWAMST2	X79907 M.musculus
42	304.4	66.0	1572	6	A44967	A44967 Sequence 23
43	304.4	66.0	1572	6	I64458	I64458 Sequence 23
44	303.2	65.8	1547	10	BC028249	BC028249 Mus muscu
45	303	65.7	747	10	MMU012555	AJ012555 Mus muscu

ALIGNMENTS

RESULT 1	AR164506	AR164506	461 bp	DNA	linear	PAT 17-OCT-2001
LOCUS	AR164506	Sequence 3 from patent US 6274143.				
DEFINITION	AR164506					
ACCESSION	AR164506					
VERSION	AR164506.1	GI:16237556				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 461)					
AUTHORS	Chatterjee,M. and Foon,K.A.					
TITLE	Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11b10					
JOURNAL	Patent: US 6274143-A 3 14-AUG-2001;					

	RESULT	3
Dd	LOCUS	AF124720
	DEFINITION	Mus musculus immunoglobulin heavy chain mRNA, partial cds.
	ACCESSION	AF124720
	VERSION	AF124720.1 GI:14164544
	KEYWORDS	
	SOURCE	Mus musculus (house mouse)
	ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 411) Tripathi,P.K., Qin,H., Bhattacharya-Chatterjee,M., Ceriani,R.L., Foon,K.A. and Chatterjee,S.K. Construction and characterization of a chimeric fusion protein consisting of an anti-idiotyp antibody mimicking a breast cancer-associated antigen and the cytokine GM-CSF Hybridoma 18 (2), 193-202 (1999)
	REFERENCE	99306687
	AUTHORS	PUBMED 10380019
	TITLE	2 (bases 1 to 411) Chatterjee,S.K. and Tripathi,P.K. Direct Submission
	JOURNAL MEDLINE	
	PUBLISHED	
	REFERENCE	
	AUTHORS	
	TITLE	Submitted (29-JAN-1999) Internal Medicine, University of Kentucky, 800 Rose Street, Lexington, KY 40536, USA
	JOURNAL	Location/Qualifiers
	FEATURES	

```
source          1. .411
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
CDS
1. .>411
/note="anti-idiotypic antibody 11D10; mimics a breast
cancer-associated antigen, human fat globule (HMPG)"
/codon_start=1
/product="immunoglobulin heavy chain"
/protein_id="AAK5119.1"
/db_xref="GI:14164545"
/translation="MCSWVFLFLSITTVHVSQAYLQSGAELVRSQVSKMSCKAS
GYTLTSYNNHVVKQTEGQLEWIGYIFLNAGTNYNOKFKGKATLTADTSSSTAYMOI
SLTSDSAVYFCARGWEGALDYWGQGTSTVSS"
BASE COUNT    102 a 101 c 107 g 101 t
ORIGIN
11. .>469
Query Match      88.8%; Score 409.4; DB 10; Length 411;
Best Local Similarity 99.8%; Pred. No. 8.5e-129;
Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGAATGAGCTGGTCTTCTCTCTCTCTCAATAACTACAGGTGCCACTCCAG 60
DB 1 ATGGAATGAGCTGGTCTTCTCTCTCTCTCAATAACTACAGGTGCCACTCCAG 60
QY 61 GCTTATCTACAGCAGCTGGGCTGAGCTGGTGGGCTGAGCTGGGCTGAGTATGCC 120
DB 61 GCTTATCTACAGCAGCTGGGCTGAGCTGGTGGGCTGAGCTGGGCTGAGTATGCC 120
QY 121 TGCAGAGCTTCTGGCTACACATTTGACCATTTACAAATATGACATGGGTAAAGCAGACCT 180
DB 121 TGCAGAGCTTCTGGCTACACATTTGACCATTTACAAATATGACATGGGTAAAGCAGACCT 180
QY 181 GGACAGGCTTGAATGGATTTGAAATATTTTCTGGAATGGTGTATCTTACTCAAT 240
DB 181 GGACAGGCTTGAATGGATTTGAAATATTTTCTGGAATGGTGTATCTTACTCAAT 240
QY 241 CAGAAGTTTAAAGGCAAGGCTCATTTGACTGCAGACACATCTCCAGCAGCCTACATG 300
DB 241 CAGAAGTTTAAAGGCAAGGCTCATTTGACTGCAGACACATCTCCAGCAGCCTACATG 300
QY 301 CAGATCAGAGCCTGACATCTGAAGACTCTGCGGTCTATTCTCTGCAAGAGGAACTGG 360
DB 301 CAGATCAGAGCCTGACATCTGAAGACTCTGCGGTCTATTCTCTGCAAGAGGAACTGG 360
QY 361 GAGGCTGCTTGACTACTGAGACTCTGAAGACTCTGAGTCAAGCACTCAGTCACTCTCTCA 411
DB 361 GAGGCTGCTTGACTACTGAGACTCTGAAGACTCTGAGTCAAGCACTCAGTCACTCTCTCA 411

RESULT 4
AB050080
LOCUS          471 bp mRNA linear ROD 02-APR-2002
DEFINITION    Mus musculus VH9H8 mRNA for anti-A/U antibody, partial cds.
ACCESSION     AB050080
VERSION       AB050080.1 GI:19909935
KEYWORDS
SOURCE        Mus musculus (house mouse)
ORGANISM
REFERENCE
1 Kitagawa, Y. and Okuhara, E.
AUTHORS
TITLE         The separation of three antibody populations from
anti-poly(A).poly(U) antibodies elicited in mice or rabbits and
antigenic features of poly(A).poly(U)
JOURNAL       Mol. Immunol. 19 (2), 257-266 (1982)
MEDLINE       82245325
PUBMED        6178956
REFERENCE
2 Kitagawa, Y.
AUTHORS
TITLE         anti-dsrNA (A/U) Ab VH region VH9H8

JOURNAL       Published Only in Database (2002)
REFERENCE     3 (bases 1 to 471)
AUTHORS       Kitagawa, Y.
TITLE         Direct Submission
JOURNAL       Submitted (17-OCT-2000) Yoshichika Kitagawa, Akita Prefectural
University, Biotechnology Institute; Minami 2-2, Ogata, Akita
010-0444, Japan (E-mail: kitagawa@agri.akita-pu.ac.jp,
URL: www.akita-pu.ac.jp, Tel: 81-185-45-2026 (ex.400),
Fax: 81-185-45-2678)
FEATURES
Location/Qualifiers
source        1. .471
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/cell_line="hybridoma 9H8"
1. .471
/gene="VH9H8"
11. .>469
/gene="VH9H8"
/codon_start=1
/product="anti-A/U antibody"
/protein_id="BAB87192.1"
/db_xref="GI:19909936"
/translation="MGWTVFLILLITTVHVSQAYLQSGAELVRSQVSKMSCKAS
GYTFTSYNNHVVKQTEGQLEWIGYIFLNAGTNYNOKFKGKATLTADTSSSTAYMOIS
SLTSDSAVYFCARGEGYGNPWFAYWGQGLTVTVSAAKTTAQPYPYPLAPGSL"
BASE COUNT    114 a 126 c 121 g 110 t
ORIGIN
77.0%; Score 355.2; DB 10; Length 471;
Best Local Similarity 88.4%; Pred. No. 3.6e-110;
Matches 410; Conservative 0; Mismatches 48; Indels 6; Gaps 2;
QY 1 ATGGAATGAGCTGGTCTTCTCTCTCTCTCAATAACTACAGGTGCCACTCCAG 60
DB 11 ATGGAATGAGCTGGTCTTCTCTCTCTCTCAATAACTACAGGTGCCACTCCAG 70
QY 61 GCTTATCTACAGCAGCTGGGCTGAGCTGGTGGGCTGAGCTGGGCTGAGTATGCC 120
DB 71 GCTTATCTACAGCAGCTGGGCTGAGCTGGTGGGCTGAGCTGGGCTGAGTATGCC 130
QY 121 TGCAGAGCTTCTGGCTACACATTTGACCATTTACAAATATGACATGGGTAAAGCAGACCT 180
DB 131 TGCAGAGCTTCTGGCTACACATTTGACCATTTACAAATATGACATGGGTAAAGCAGACCT 190
QY 181 GGACAGGCTTGAATGGATTTGAAATATTTTCTGGAATGGTGTATCTTACTCAAT 240
DB 191 GGACAGGCTTGAATGGATTTGAAATATTTTCTGGAATGGTGTATCTTACTCAAT 247
QY 241 CAGAAGTTTAAAGGCAAGGCTCATTTGACTGCAGACACATCTCCAGCAGCCTACATG 300
DB 248 CAGAAGTTTAAAGGCAAGGCTCATTTGACTGCAGACACATCTCCAGCAGCCTACATG 307
QY 301 CAGATCAGAGCCTGACATCTGAAGACTCTGCGGTCTATTCTCTGCAAGAGGAACT-- 358
DB 308 CAGATCAGAGCCTGACATCTGAAGACTCTGCGGTCTATTCTCTGCAAGAGGAGCTAT 367
QY 359 -GGAGGGTCTCTGGACTACTGGGTCAAGGAACTCAGTCACTCCTCTCTCAGCAAA 417
DB 368 GGTAAACCCCTCTGTTGCTTACTGGGCCAAGGAACTCTGGTCACTCTCTCTGAGCCAAA 427
QY 418 ACCACACCCCCCTCTGTTGCTTACTGGGCCAAGGAACTCTGGTCACTCTCTCTGAGCCAAA 461
DB 428 ACAACAGCCCAACCCCTTATCCCTTGGCCCTTGGAGCTTGGG 471

RESULT 5
AY178830
LOCUS          457 bp mRNA linear ROD 15-JAN-2003
DEFINITION    Mus musculus clone 13G10 antiporphyrin immunoglobulin G heavy chain
variable region mRNA, partial cds.
ACCESSION     AY178830
```



gene

```
1. .1542
/ gene="IGH-4"
/ note="synonym: IgG1"
/ db_xref="LocusID:16017"
/ db_xref="MGI:96446"
33. .1418
/ codon_start=1
/ product="Igh-4 protein"
/ protein_id="AAH02121.1"
/ db_xref="GI:12805309"
/ db_xref="LocusID:16017"
/ translation="NEWIWFILSLTASVHSQVLOQSGAELARPAGSVKMSKAS
GYSFYSGLSWKQTRQGLEWIGIYFSGNYYNEFKGKATLATADKSSATYMQ
SLTSEDYAVFCAVWGPWDYWGQGTITVSSAKTTPSVYPLAPGSAATNSMV
TLGCLVKGYFPEPTVTWNSGLSSGVHTFPALQSDLYTLSSVTPSPSTQVTV
CNVAHPASKVDKIVPRDCGCKPCICTVPEVSSVFIIPPKPKDLVLTILPKVTCV
VVDISKDDPEQVFSFVDDVEVHTAQTPREOFNSTRSVSELPIMHQDWLNGKEFK
CRVNSAAPAPLEKIYISIKGRPKAPQVITIPPREQNAKOKVSLTMITDFPEDIT
VEWQNGQPAENYKNTQPIMDTGDGSFYVLSKLVQKSNWEAGNTFTCSVLHGLNHH
TEKLSLSHSPGK"
BASE COUNT 406 a 447 c 364 g 325 t
ORIGIN
```

```
Query Match 73.3%; Score 337.8; DB 10; Length 1542;
Best Local Similarity 84.1%; Pred. No. 4.2e-104;
Matches 381; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 ATGGAATGAGCTGGGCTCTTCTCTCCCTGTCATTAATACAGGTGTCACATCCCAG 60
DB 33 ATGGAATGAGCTGGGCTCTTCTCTCCCTGTCATTAATACAGGTGTCACATCCCAG 92
QY 61 GCTTATCTACAGCAGTCTGGGCTGAGCTGGTGGTGGGCTCTGAGGATGTCC 120
DB 93 GTTCAGCTGCAGCAGTCTGGGCTGAGCTGGGCTGGGCTCTGAGGATGTCC 152
QY 121 TCAGAGCTTCTGGCTACACATGACGATTAACATATGACATGCGTAAAGCAGACACCT 180
DB 153 TCGAAGGCTTCTGGTACTCATTCACAGCTATGATTAAGTGGTGAAGCAGAAACT 212
QY 181 GGACAGGCTTGAATGAGTGAATATTTTCTGGAAATGGTATCTACTACAAAT 240
DB 213 AGACAGGCTTGAATGAGTGAATATTTTCTGGAAATGGTATCTACTACAAAT 272
QY 241 CAGAAGTTAAGGCAAGGCTTATGACTGCAGACACATCTCCAGCAGACGCTACATG 300
DB 273 GAAAAGTTCAAGGCAAGGCTTATGACTGCAGACACATCTCCAGCAGACGCTACATG 332
QY 301 CAGATCAGCAGCTGACATCTGAAGCTCTGGGCTTATTTCTGTCAGAGAGGAGCTGG 360
DB 333 CAGCTCAGCAGCTGACATCTGAGGCTCTGAGGCTTATTTCTGTCAGAGATGGGGGAG 392
QY 361 GAGGCTCTCTGAGTCTGAGGCTCAAGGAACTCAGTCAAGCTCTCTCTGTCAGAGAGG 420
DB 393 CCCTGGGAGCTGAGTCTGAGGCTCAAGGAACTCAGTCAAGCTCTCTCTGTCAGAGAGG 452
QY 421 ACACCCCAAGCTCTATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 453
DB 453 ACACCCCAAGCTCTATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 485
```

RESULT 7

```
MMU39899 483 bp DNA linear ROD 21-NOV-1996
LOCUS Mus musculus anti-glycoprotein-B of human Cytomegalovirus
DEFINITION Immunoglobulin Vh chain gene, partial cds.
ACCESSION U39899
VERSION U39899.1 GI:1680666
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 483)
```

AUTHORS

Schoppel, K., Hassfurther, E., Britt, W., Ohlin, M., Borrebaeck, C.A. and Mach, M.  
Antibodies specific for the antigenic domain 1 of glycoprotein B (gpUL55) of human cytomegalovirus bind to different substructures Virology 216 (1), 133-145 (1996)  
96187797  
MEDLINE  
PUBMED 8614980  
REFERENCE 2 (bases 1 to 483)  
AUTHORS Schoppel, K.  
TITLE Direct Submission  
JOURNAL Submitted (02-NOV-1995) Michael Mach, Institute of Virology, University Erlangen-Nuernberg, Schlossgarten 4., 91054 Erlangen, Germany

FEATURES

source  
1. .483  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="Balb/C"  
/db\_xref="taxon:10090"  
/note="monoclonal antibody 27-156"  
1. .>483  
/codon\_start=1  
/product="anti-glycoprotein-B of human Cytomegalovirus Immunoglobulin Vh chain"  
/protein\_id="AAB26952.1"  
/db\_xref="GI:1680667"

CDS

```
BASE COUNT 117 a 118 c 130 g 118 t
ORIGIN
Query Match 71.2%; Score 328.2; DB 10; Length 483;
Best Local Similarity 83.8%; Pred. No. 6.9e-101;
Matches 388; Conservative 0; Mismatches 63; Indels 12; Gaps 1;

QY 1 ATGGAATGAGCTGGGCTCTTCTCTCCCTGTCATTAATACAGGTGTCACATCCCAG 60
DB 1 ATGGAATGAGCTGGGCTCTTCTCTCCCTGTCATTAATACAGGTGTCACATCCCAG 60
QY 61 GCTTATCTACAGCAGTCTGGGCTGAGCTGGTGGTGGGCTCTGAGGATGTCC 120
DB 61 GTTCAGCTGCAGCAGTCTGGGCTGAGCTGGTGGTGGGCTCTGAGGATGTCC 120
QY 121 TCGAAGGCTTCTGGCTTACACATTCACAGTTCACATATGACATGCGTAAAGCAGACACCT 180
DB 121 TCGAAGGCTTCTGGCTTACACATTCACAGTTCACATATGAGTGGTGAAGCAGGCT 180
QY 181 GGACAGGCTTGAATGAGTGAATATTTTCTGGAAATGGTATCTACTACAAAT 240
DB 181 GGACAGGCTTGAATGAGTGAATATTTTCTGGAAATGGTATCTACTACAAAT 240
QY 241 CAGAAGTTAAGGCAAGGCTTATGACTGCAGACACATCTCCAGCAGACGCTACATG 300
DB 241 GAAAAGTTCAAGGCAAGGCTTATGACTGCAGACACATCTCCAGCAGACGCTACATG 300
QY 301 CAGATCAGCAGCTGACATCTGAAGCTCTGGGCTTATTTCTGTCAGAGAGGAACTGG 360
DB 301 CAGCTCAGCAGCTTACATCTGAGGCTCTGGGCTTATTTCTGTCAGAGAGGTTACGAC 360
QY 361 GAGG-----GTGCTCTGAGTCTGAGGCTCAAGGAACTCTACTGTCACGCTGCTCC 408
DB 361 GAAGTGGATTTACTTTGCTATGGACTACTGGGCTCAAGGAACTCTACTGTCACGCTCC 420
QY 409 TCAGCCAAACACACACCCCTCTATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451
DB 421 TCAGCCAAACACACACCCCTCTATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 463
```

RESULT 8

BD092129 409 bp DNA linear PAT 27-AUG-2002  
LOCUS

KEYWORDS	Mus musculus (house mouse)
SOURCE ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE AUTHORS	1 (bases 1 to 1544) Takahashi, S., Itoh, J., Nose, M., Ono, M., Yamamoto, T. and Kyogoku, M.
TITLE	Cloning and cDNA sequence analysis of nephritogenic monoclonal antibodies derived from an MRL/lpr lupus mouse
JOURNAL MEDLINE PUBMED	Mol. Immunol. 30 (2), 177-182 (1993) 93156722
REFERENCE	8429833
AUTHORS	2 Ono, M., Yamamoto, T., Kyogoku, M. and Nose, M.
TITLE	Sequence analysis of the germ-line VH gene corresponding to a nephritogenic antibody in MRL/lpr lupus mice
JOURNAL MEDLINE PUBMED	Clin. Exp. Immunol. 100 (2), 284-290 (1995) 95262286
REFERENCE	7743668
AUTHORS	3 (bases 1 to 1544) Nose, M.
TITLE	Direct Submission
JOURNAL	Submitted (10-MAR-1993) Masato Nose, Tohoku University School of Medicine, Dept. of Pathology, 2-1 Seiryō-cho Aoba-ku, Sendai, Miyagi 980, Japan (E-mail:d32181@cctu.cc.tohoku.ac.jp,
FEATURES source	Tel:81-22-273-9042, Fax:81-22-234-1986) Location/Qualifiers 1. .1544 /organism="Mus musculus" /mol_type="mRNA" /strain="MRL/MpJ-lpr/lpr" /db_xref="taxon:10090"

```

/chrosome="1"
/clone="PB1H1"
/clone_lib="pBluescript KS(+)"
/dev_stage="adult"
1. .32
33. .1445
/note="precursor of C and V-D-J regions from 7B6.8"
/codon_start=1
/product="immunoglobulin gamma-3 heavy chain precursor"
/protein_id="BAA03476.1"
/db_xref="GI:1304160"
/translation="MEWSGVFFLLSVTAGVHSQVLOQSGAELVRPGTSVRKMSCKAA
GTYFTNWIGWKVQRPGHGIWDIYPGGGYNYNEKPKGKATLADTSSSTAYMQLG
SSLTSDSAIYCYCARLGGYGYVAMDYWGQTSVTSSATTTAPSVYPLVPGSCDSTGS
SWTGLCLVKYGVPEPVTWKMYGALSGVRTVSSVLSQGFYSLISLVTVPSSWTPSQ
TVICNVAHSPAKTELKRIEPRIPKESTPGSCCPGNI LGGSVFIFPPKPKDRLMI
SLTPKVTCCVVDSEDDPDVFNVDNKEVHTAWTPQRAQYNSTFRVVSALPIQHLQ
DMWRGKFKVNNKALPAIERITSPKGRAGTPQVYIPPPREQMSKKKKVSLTCLV
TNFFSAISGVERNELGEQDKYNTPTPILDSQGTIFYLYSKLTVDTDSWLQGEITFGVS
VREALHNHHTQNLRSRSPGK"
33. .89
90. .1442
/product="immunoglobulin gamma-3 heavy chain"
90. .452
/product="immunoglobulin V-D-J region"
/note="7B6.8"
453. .1442
/product="immunoglobulin gamma-3 secreted C-region"
1446. .1544
1514. .1519
BASE COUNT 402 a 448 c 368 g 326 t
3'UTR
polyA_signal
ORIGIN

```

	Query Match	89.0%;	Score	3267; DB	60;	Length	1354;
	Best Local Similarity	83.0%;	Pred.	No. 1.1e-99;			
	Matches	386;	Conservative	0;	Mismatches	70;	Indels
	Gaps	9;					
QY	1	ATGGAATGCAGCTGGTCTTTTCCTCCCTCAATAACTACAGTGTCCTACTCCCAG	60				
DZ	33	ATTGGAATGGAGGGGGGTCTTTATCTTCTCAGTAAGTGCAGTGTCCTACTCCCAG	92				

61	QY	GC	TAT	TCTACAGCAGTCTGGGCTGACCTGGT	GAGGTC	TCTGGGGCCTCAGTGAAGATGCC	120
93	Db	GT	CCAACTTCAGCAGCTCTGGAGCTGAGCT	GAGCTGGT	TAGGGCTGGGAGCTTCAGTGAAGATGCC	152	
121	QY	TG	CAAGGCTCTCGCTACACATTCACCA	TTACACATATGCAC	TGGGTAAAGCAGACACACT	180	
153	Db	TG	CAAGGCTGCTGGATACACCTTCAC	TAACTAC	TGGATAGGTTGGGTAAAGCAGAGGCT	212	
181	QY	GG	ACAGGGCTGGAATGGATTGGAAATAT	TTTTCTCGAAATGGTGATACTTACTACAAT		240	
213	Db	GG	ACATGGCCTTCAGTGGATTGGAGATAT	TATACCTGGAGTGTTTATACTAATACTACAAT		272	
241	QY	CAGA	AGTTTAAAGGCAAGGCTCAT	TGACTGCAGACACATCTCCAGCAGACGCTACATG		300	
273	Db	GAGAAATTC	CAAGGGCAAGGCCACACTGACT	TCAGACACATCTCCAGCAGACGCTACATG		332	
301	QY	CAGATCAG	CAGCTGCATCTGGAAGACTCTCGG	GTCTATTCTGTGCAAGGAAGCACTGG		360	
333	Db	CAC	TCAAGCCTGCATCTGAGGACTCTGCC	ATCTATTACTGTGCAAGACTGAGGGGG		392	
361	QY	GAGGG-----TG	CTTGGACTACTGGGGTCAAGGA	ACCTCAGTCAACGCTCTCCTCA		411	
393	Db	GGGGGATAT	TACTATGCGACTACTGGGTCAAG	GAACCTCAGTCAACGCTCTCCTCA		452	
412	QY	GC	CAAAACGACACCCCAACCGTCTAT	TCAC	TGGTCCCTGGAAGC	456	
453	Db	GC	TACAAACACACCCCATCTGCTAT	CCCTTGGTCCCTGGCTGC		497	

[illegible]

QY	301	CAGATCAGCAGCCTGACATCTGAAGACTCTGCGGCTATTTCGTGCAAGAGGAACTGG	360
Db	301	CAGCTCAGCAGCCTGACATCTGAGGACTCTGCGCATCTATTACTTGTCAAGGGGGATACCG	360
QY	361	GAGGCTGCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCTCAGGCCAAAAG	420
Db	361	GGATATGCTATGGACTTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCTCAGGCCAAAACA	420
QY	421	ACACCCCCACCCGCTCTATCCACT	443
Db	421	ACAGCCCCCATCGGCTCTATGCAC	443

RESULT 11	
MUSIGHALPA	
LOCUS	525 bp mRNA linear ROD 18-NOV-1994
DEFINITION	Mus musculus IgG2a chain (anti-Pseudomonas aeruginosa lipoprotein I antibody) mRNA, 5' end.
ACCESSION	M28834
VERSION	M28834.1 GI:576597
KEYWORDS	V-segment; anti-lipoprotein antibody; immunoglobulin G2a gamma chain; immunoglobulin heavy chain.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; 1 (bases 1 to 525)
AUTHORS	Margel,M., Eckhardt,A., Ehret,W., von Specht,B.U., Duchene,M. and Domdey,H.
TITLE	Cloning and characterization of cDNAs coding for the heavy and light chains of a monoclonal antibody specific for Pseudomonas aeruginosa outer membrane protein I
JOURNAL	Gene 74 (2), 335-345 (1988)
MEDLINE	89232725
PUBMED	3149944
COMMENT	On Nov 28, 1994 this sequence version replaced gi:342018. Original source text: Mus musculus (strain BALB/c, sub_species domesticus) hybridoma cDNA to mRNA.

```

gene      /lssoc_type= nybriuoma
1. .525
/gene="Igg"
CDS
49. .>525
/gene="Igg"
/codon_start=1
/product="immunoglobulin gamma-2a chain"
/protein_id="AA53291.1"
/db_xref="GI:576598"
/translation="MERHWIFLFESVTAGVHSQVLOQSGAELAKPGASVKNMCKAS
GYTFAYMHWVKORPGQGLEWIGYNPNTGYEYNQNFKDATLTADKSSSTAYMQL
SSLTSDSAVYCYTRSYNYNVEGAMDYWGQGTSTVTVSSAKTTAPSYVPLAPVCGDITG"
49. .105
/sig_peptide
/gene="Igg"
mat_peptide
106. .>525
/gene="Igg"
/product="immunoglobulin gamma-2a chain"

BASE COUNT      144 a  142 c  119 g  120 t
ORIGIN

Query Match      70.1%; Score 323; DB 10; Length 525;
Best Local Similarity 83.4%; Pred. No. 4.3e-99;
Matches 381; Conservative 0; Mismatches 70; Indels 6; Gaps 1;

Qy      1  ATGGAATGCAGCTGGGCTTTCTTCCTCCCTGTCATAACTACAGGTGTCCACTCCCGAG 60
      ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |

```



```

Db      49  ATGGAAGGCACTGGATCTTCTCTCTGTTTTCAGTTACTGCGAGGTGTCACACTCCAG 108
QY      61  GCTTATCTACAGCACTCTGGGCTGAGCTGGTGTGAGGCTCTGGGCTCAGTGAAGATGCC 120
Db      109  GTCCAGCTTACAGCACTCTGGGCTGAACTGGCAAAACCTGGGCTCAGTGAAGATGCC 168
QY      121  TGCAGGCTTCTGGCTACACATTCACCACTTACATATGCACTGGTAAACAGACACCT 180
Db      169  TCGAAGGCTTCTGGCTACACCTTACTGCTTACTGCTACTGGATGCACTGGTAAACAGAGGCT 228
QY      181  GCAGAGGCTTGGATGATGGAATATTTTCTGGAATGTTGATGATCTACTACAAT 240
Db      229  GCAGAGGCTTGGATGATGGAATATTTTCTGGAATGTTGATGATCTACTACAAT 288
QY      241  CAGAAGTTTAAGGCAAGGCTCATTTGACTGCAGACACATCTCCAGCACAGCCCTACATG 300
Db      289  CAGAATTTCAAGGCAAGGCTCATTTGACTGCAGACAAATCTCCAGCACAGCCCTACATG 348
QY      301  CAGATCAGCAGCTGACATCTGAGACTCTGCGGTCTATTCTGTGCAAGAGG-----G 354
Db      349  CAACTGAGCAGCTGACATCTGAGGACTCTGAGTCTATTATTGTACAAGAGCTACTAT 408
QY      355  AACTGGGAGGTGCTCTGGACTACTGGGCTCAAGGAACCTCAGTCACCGTCTCTCTCAGCC 414
Db      409  AACTAGAGGGGCTATGAGTACTGGGCTCAAGGAACCTCAGTCACCGTCTCTCTCAGCC 468
QY      415  AAAACGACACCCACCGCTCTATCCACTGGTCCCTG 451
Db      469  AAAACAACAGCCCCATCGTCTATCCACTGGCCCTG 505

```

## RESULT 12

Al3735 540 bp DNA linear PAT 03-JAN-1994  
 DEFINITION variable region of a monoclonal antibody which cross reacts with 19 known Pseudomonas aeruginosa serotypes.

ACCESSION Al3735

VERSION Al3735.1 GI:491743

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1 (bases 1 to 540)

AUTHORS Domdey,H., Margel,M. and von Specht,B.U.

TITLE Monoclonal antibodies to Pseudomonas aeruginosa, their production and use

JOURNAL Patent: EP 0338395-A 3 25-OCT-1989;

BEHRINGWERKE Aktiengesellschaft

FEATURES Location/Qualifiers

source 1..540

/organism="synthetic construct"

/mol\_type="genomic DNA"

/db\_xref="taxon:32630"

64..&gt;540

/codon\_start=1

/transl\_table=1

/product="v region monoclonal antibody,cross reacts with

19 known P.aeruginosa serotypes"

/protein\_id="CAA01127.1"

/db\_xref="GI:491744"

/translation="MERWIFLFLFSVTAGVHSQVLOQSGAELAKPGASVKMSCKAS

GITFTAWHWKRPQGLWIGTINPTGYTNEONFKDKATILADKSSSTAYMQL

SSLTSEDSAVYITRSYNYEGADYNGQSTVTSVSSAKTTAPSVYPLAPVCGDTTG"

BASE COUNT 145 a 144 c 130 g 121 t

ORIGIN

Query Match 70.1%; Score 323; DB 6; Length 540;

Best Local Similarity 83.4%; Pred. No. 4.3e-99;

Matches 381; Conservative 0; Mismatches 70; Indels 6; Gaps 1;

QY 1 ATGGAATCAGCTGGTCTTCTCTCTGTTTTCAGTTACTGCGAGGTGTCACACTCCAG 60

Db 64 ATGGAAGGCACTGGATCTTCTCTCTGTTTTCAGTTACTGCGAGGTGTCACACTCCAG 123

```

QY      61  GCTTATCTACAGCACTCTGGGCTGAGCTGGTGTGAGGCTCTGGGCTCAGTGAAGATGCC 120
Db      124  GTCCAGCTTACAGCACTCTGGGCTGAACTGGCAAAACCTGGGCTCAGTGAAGATGCC 183
QY      121  TGCAGGCTTCTGGCTACACATTCACCACTTACATATGCACTGGTAAACAGACACCT 180
Db      184  TCGAAGGCTTCTGGCTACACCTTACTGCTTACTGCTTACTGGATGCACTGGTAAACAGAGGCT 243
QY      181  GCAGAGGCTTGGATGATGGAATATTTTCTGGAATGTTGATGATCTACTACAAT 240
Db      244  GCAGAGGCTTGGATGATGGAATATTTTCTGGAATGTTGATGATCTACTACAAT 303
QY      241  CAGAAGTTTAAGGCAAGGCTCATTTGACTGCAGACACATCTCCAGCACAGCCCTACATG 300
Db      304  CAGAATTTCAAGGCAAGGCTCATTTGACTGCAGACAAATCTCCAGCACAGCCCTACATG 363
QY      301  CAGATCAGCAGCTGACATCTGAGACTCTGCGGTCTATTCTGTGCAAGAGG-----G 354
Db      364  CAACTGAGCAGCTGACATCTGAGGACTCTGAGTCTATTATTGTACAAGAGCTACTAT 423
QY      355  AACTGGGAGGTGCTCTGGACTACTGGGCTCAAGGAACCTCAGTCACCGTCTCTCTCAGCC 414
Db      424  AACTAGAGGGGCTATGAGTACTGGGCTCAAGGAACCTCAGTCACCGTCTCTCTCAGCC 483
QY      415  AAAACGACACCCACCGCTCTATCCACTGGTCCCTG 451
Db      484  AAAACAACAGCCCCATCGTCTATCCACTGGCCCTG 520

```

## RESULT 13

I05921

LOCUS I05921

DEFINITION Sequence 37 from Patent EP 0274394.

ACCESSION I05921

VERSION I05921.1 GI:590876

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 458)

AUTHORS Robinson,R.R., Liu,A.Y., Hellstrom,K.E., Hellstrom,I. and

Ledbetter,J.A.

TITLE Chimeric antibody with specificity to human B cell surface antigen

JOURNAL Patent: EP 0274394-A2 37 13-JUL-1988;

FEATURES Location/Qualifiers

source 1..458

BASE COUNT 113 a 120 c 112 g 113 t

ORIGIN

Query Match 69.8%; Score 322; DB 6; Length 458;

Best Local Similarity 94.4%; Pred. No. 9.2e-99;

Matches 334; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGGAATCAGCTGGTCTTCTCTCTGTTTTCAGTTACTGCGAGGTGTCACACTCCAG 60

Db 39 ATGGGATTCACAGCACTCTTCTCTCTGTTTTCAGTTACTGCGAGGTGTCACACTCCAG 98

QY 61 GCTTATCTACAGCACTGGGCTGAGCTGGTGTGAGGCTCTGGGCTCAGTGAAGATGCC 120

Db 99 GCTTATCTACAGCACTGGGCTGAGCTGGTGTGAGGCTCTGGGCTCAGTGAAGATGCC 158

QY 121 TGCAGGCTTCTGGCTACACATTCACCACTTACATATGCACTGGGTAAGACAGACCT 180

Db 159 TCGAAGGCTTCTGGCTACACATTCACCACTTACATATGCACTGGGTAAGACAGACCT 218

QY 181 GCAGAGGCTTGGATGATGGAATATTTTCTGGAATGTTGATGATCTACTACAAT 240

Db 219 AGACAGGCTTGGATGATGGAATATTTTCTGGAATGTTGATGATCTACTACAAT 278

QY 241 CAGAAGTTTAAGGCAAGGCTCATTTGACTGCAGACACATCTCCAGCACAGCCCTACATG 300





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 16:05:49 ; Search time 202.202 Seconds  
(without alignments)  
6154.444 Million cell updates/sec

Title: us-08-836-455-3  
Perfect score: 461  
Sequence: 1 ATGGATCAGCTGGCTCTT.....CTGGTCCCTGGAGCTTGGG 461

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
N_Geneseq_19Jun03.*			
1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*		
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*		
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*		
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*		
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*		
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*		
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*		
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*		
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*		
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*		
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*		
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*		
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*		
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*		
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*		
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*		
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*		
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*		
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*		
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*		
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*		
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*		
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*		
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*		
25:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	461	100.0	461	18	AAT85150 Murine monoclonal
2	461	100.0	461	20	AAV83773 Antibody 11D10 hea
3	461	100.0	461	25	AAV51274 Mouse 11D10 antibo
4	327.2	71.0	409	21	AAZ87715 Anti-human VEGF re
5	327.2	71.0	409	22	AAV70192 Anti-human Flt-1 m
6	326.6	70.8	736	25	ABX16571 Mouse DNA encoding
7	323	70.1	540	10	AAV91645 Heavy chain of mon
8	322	69.8	458	10	AAV91146 2H7 Vh sequence.

9	322	69.8	459	20	AAV82357 Mouse antibody 2H7
10	322	69.8	459	22	AAH22070 2H7 heavy chain va
11	322	69.8	491	18	AAT70868 Mouse 2H7 antibody
12	322	69.8	491	19	AAV03926 Mouse 2H7 antibody
13	322	69.8	491	19	AAV03926 Mouse 2H7 antibody
14	322	69.8	491	19	AAV18593 Mouse 2H7 antibody
15	322	69.8	518	18	AAT36316 Coding sequence to
16	322	69.8	520	18	AAT51042 Anti-tobacco mosai
17	320.6	69.5	1553	16	AAV79930 Mouse immunoglobul
18	320	69.4	1683	20	AAV61084 Consensus DNA sequ
19	314.8	68.3	494	19	AAV20085 Monoclonal antibod
20	311.6	67.6	1570	12	AAQ12637 Monoclonal antibod
21	310.2	67.3	1581	14	AAQ48037 2H7 Vh sequence wh
22	309.4	67.1	459	8	AAV70971 Anti-CA125 bifunct
23	308.6	66.9	420	22	AAV81910 H-chain V-region o
24	308.2	66.9	588	14	AAQ43385 3F4 (Chimeric) hum
25	308.2	66.9	1392	18	AAT62936 Murine anti-porcine
26	308.2	66.9	1395	18	AAT62935 3F4 human IgG4 ch
27	308.2	66.9	3400	18	AAT62937 3F4 human IgG4 exp
28	308.2	66.9	5300	18	AAT62938 Nucleotide sequenc
29	306.4	66.5	468	19	AAV05544 Mouse secreted exp
30	305	66.2	725	21	AAA43821 MAb 55.1 heavy cha
31	304.4	66.0	1582	16	AAQ94037 DNA encoding the h
32	299.8	65.0	440	19	AAV22074 VH186 region of an
33	299	64.9	458	12	AAQ15164 H chain subunit of
34	297.4	64.5	1773	18	AAT88869 Anti-human Fas mon
35	297.4	64.5	1773	19	AAV66735 Anti-human Fas ant
36	297.4	64.5	1773	21	AAV78202 Human CH2 domain d
37	297.2	64.5	1083	24	AAV45753 Antibody C2B8 CH2
38	297.2	64.5	1083	25	ABZ24017 DNA encoding a dim
39	297.2	64.5	1413	21	AAV63531 Human C2B8 antibod
40	297.2	64.5	1413	24	AAV45752 Antibody C2B8 heav
41	297.2	64.5	1413	25	ABZ24016 Vector contg. TCAE
42	297.2	64.5	9208	15	AAQ65629 TCAE8 expression v
43	297.2	64.5	9209	25	ABX94203 Traget plasmid Mol
44	297.2	64.5	19001	19	AAV61793 Plasmid pAH4808.
45	297	64.4	10844	14	AAQ43848

ALIGNMENTS

RESULT 1  
AAT85150  
ID AAT85150 standard; cDNA; 461 BP.

XX	AC	AAT85150;			
XX	AC	AAT85150;			
XX	DT	25-MAR-2003 (updated)			
XX	DT	04-JAN-1998 (first entry)			
XX	DE	Murine monoclonal anti-idiotyp antibody 11D10 Vh cDNA.			
XX	DE	Monoclonal antibody 11D10; anti-idiotyp antibody; mucin;			
XX	KW	human milk fat globule; HMG; tumour; breast cancer; vaccine; ss.			
XX	XX	Mus musculus.			
XX	XX	OS			
XX	XX	Key	Location/Qualifiers		
XX	XX	sig_peptide	1..57		
XX	XX	mat_peptide	/*tag= a		
XX	XX	mat_peptide	58..461		
XX	XX	mat_peptide	/*tag= b		
XX	XX	WO9722699-A2.			
XX	XX	26-JUN-1997.			
XX	XX	19-DEC-1996;	96WO-US20757.		
XX	XX	20-DEC-1995;	95US-0575762.		
XX	XX	26-JAN-1996;	96US-0591965.		
XX	XX	13-DEC-1996;	96US-0766350.		





PT	Gene recombinant antibodies, useful for diagnosis and as remedies for diseases due to abnormal neovascularization e.g. proliferation or metastasis of solid tumor, rheumatoid arthritis, diabetic retinopathy and psoriasis -
XX	-
PS	Disclosure; Page 155-156; 210pp; Japanese.
CC	The invention relates to a gene recombinant antibody that has specific reaction with human vascular endothelial growth factor (VEGF) receptor Flt-1. The antibodies are useful for diagnosis and as remedies for diseases due to abnormal neovascularisation such as proliferation or metastasis of solid tumor, arthritis in rheumatoid arthritis, diabetic retinopathy, premature retinopathy and psoriasis.
XX	
SQ	Sequence 409 BP; 99 A; 105 C; 99 G; 106 T; 0 other;
Query Match	71.0%; Score 327.2; DB 21; Length 409;
Best Local Similarity	88.8%; Pred. No. 1e-80;
Matches 366; Conservative	0; Mismatches 43; Indels 3; Gaps 1;
QY	1 ATGGAATGCAGCTGGGTCTTCTCCTCCTCTCAATACTACAGGTGCCACTCCCAG 60 
DB	1 ATGGGATTTCAGCAGGATCTTCTCTCCTCTGTCAGTGACTCAGGTGCCACTCCCAG 60 
QY	61 GCATTATCTACAGCAGCTCGGGCTGAGCTGGTAGAGCTCTGGGGCTCAGTGAAGATGTCC 120 
DB	61 GCITTTCTACAGCAGCTCGGGCTGAGCTGGTAGAGCTCTGGGGCTCAGTGAAGATGTCC 120 
QY	121 TGCAAGCTTCTGGCTACACATTGACCAGTTACAATATGCAGCTGGGTAAACGACACCT 180 
DB	121 TGCAAGCTTCTGGCTACACATTATCAATATACAATATGCAGCTGGGTAAACGACACCT 180 
QY	181 GGACAGGGCTGGAATGGATTGGAAATATTTTCTCTGGAATGTGTACTTACTACAAT 240 
DB	181 AGACAGGGCTGGAATGGANTTGGAGCTATTTTCCAGGAATGGTTTTACTTCTACAAT 240 
QY	241 CAGAAGTTTAAGGGCAAGGCTCATTTGACTGCAGACACATCTCCAGCACAGCCTACATG 300 
DB	241 CAGAAGTTCAAGGGCAAGGCCACACTGACTGTAGACAAAATCCTCCAGCACAGTCTACATG 300 
QY	301 CAGATCAGCAGCCTGCATCTGAAGACTCTGCGGTCTATTTCTGTGCAAGAGGAACCTGG 360 
DB	301 CAGCTCGCAGCCTGCATCTGAAGACTCTGCGGTCTATTTCTGTGCAAGAG---ATGGT 357 
QY	361 GAGGGTGCTCTGGACTACTGGGGTCAAGGAACCTCAGTCAACCGTCTCCTCAG 412 
DB	358 GACTATTACTTTGACTACTGGGGCCAAGGCCACCACCTCTCAGACTCTCCTCAG 409 
RESULT 5	
AAF70192	
ID	AAF70192 standard; DNA; 409 BP.
XX	
AC	AAF70192;
XX	
DT	20-APR-2001 (first entry)
XX	
DE	Anti-human Flt-1 monoclonal antibody KM1750 H chain V region DNA.
XX	
KW	Differentiation; monocyte; macrophage; haematopoietic stem cell; cancer;
KW	vascular endothelial growth factor; VEGF; Flt-1; inflammation; antibody;
KW	delayed hypersensitivity; malignant tumour; arteriosclerosis; ds.
XX	
OS	Mus musculus.
XX	
PX	WO200079275-A1.
NN	
PD	28-DEC-2000.
XX	
PF	16-JUN-2000; 2000WO-JP03957.
XX	
PR	17-JUN-1999; 99JP-0171709.
XX	











```
PR 01-NOV-1985; 85US-0793980.
PR 27-OCT-1986; 86WO-US02269.
PR 24-JUL-1987; 87US-0077528.
PR 11-JAN-1988; 88US-0142039.
PR 08-DEC-1992; 92US-0987555.
PR 18-AUG-1994; 94US-0299085.
PR 06-JUN-1995; 95US-0467140.
XX
XX (XOMA ) XOMA CORP.
XX
XX Better M, Horwitz AH, Lei S, Liu AY, Robinson RR;
XX Wall R, Wilcox GL;
XX
XX WPI; 1998-051492/05.
XX P-PSDB; AAW47513.
XX
XX DNA encoding secretable immunoglobulin fragments - comprising at
XX least the variable regions of light or heavy chains
XX
XX Example IV; Fig 21; 98pp; English.
XX
XX The present sequence was used in the development of a novel method
XX for the production of an immunoglobulin (Ig) fragment capable of
XX binding an antigen. The method comprises culturing an E. coli host
XX that has been transformed with a nucleic acid molecule encoding the
XX Ig fragment, under conditions so that the Ig fragment is produced
XX and secreted. The nucleic acid molecule comprises DNA sequences
XX encoding: (a) pectate lyase secretion signal sequence operably
XX linked to a DNA sequence encoding at least the variable region of
XX an Ig Fd molecule; and (b) pectate lyase secretion signal sequence
XX operably linked to a DNA sequence encoding at least the variable
XX region of an Ig light chain, where (a) and (b) are operably linked
XX to a single prokaryotic promoter to form a dicistronic
XX transcription unit. The method is used to produce chimeric Fab
XX molecules, e.g. derived from murine monoclonal antibody 2H7 raised
XX against human B-cell surface antigen. The invention provides a
XX novel approach for producing genetically engineered antibodies of
XX desired variable region specificity and constant region
XX properties. The cloned Ig gene products can be produced by
XX expression in genetically engineered organisms. The application of
XX chemical gene synthesis, recombinant DNA cloning and production of
XX specific Ig chains in various organisms provides an effective
XX solution for the efficient large scale production of human
XX monoclonal antibodies. The invention also provides a solution to
XX the problem of class switching antibody molecules.
XX
XX Sequence 491 BP; 113 A; 152 C; 113 G; 113 T; 0 other;
XX
XX Query Match 69.8%; Score 322; DB 19; Length 491;
XX Best Local Similarity 94.4%; Pred. No. 2.9e-79;
XX Matches 334; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
XX
XX QY 1 ATGGATGAGCTGGGCTCTTCTCTCCCTGTCATTAACACAGGTGTCACCTCCAG 60
XX DB 72 ATGGGATTCAGCAGGATCTTCTCTCCCTGTCAGTAACTACAGGTGTCACCTCCAG 131
XX QY 61 GCTTATCTACAGCAGTCTGGGCTGAGCTGGTGAGCTCTGGGCTCAGTGAAGATGTC 120
XX DB 132 GCTTATCTACAGCAGTCTGGGCTGAGCTGGTGAGCTCTGGGCTCAGTGAAGATGTC 191
XX QY 121 TGAAGGCTTCTGGCTACACATTTGACCATTTACATATGACATGCTGGTAAAGCAGACCT 180
XX DB 192 TGAAGGCTTCTGGCTACACATTTACCATTTACATATGACATGCTGGTAAAGCAGACCT 251
XX QY 181 GGACGGGCTGGAATGGATGGAAATATTTCTGGAAATGGTATCTACTACTCAAT 240
XX DB 252 AGACAGGGCTGGAATGGATGGAGCTATTTATCCAGGAATGGTATCTACTACTCAAT 311
XX QY 241 CAGAAGTTAAAGGCAAGGCTCATTTGACTGCGACACACATCTCCACAGCAGCTACATG 300
XX DB 312 CAGAAGTTAAAGGCAAGGCTCATTTGACTGCGACACACATCTCCACAGCAGCTACATG 371
XX QY 301 CAGATCAGCAGCTGACATCTGAAGACTCTGCGGCTCTATTTCTGTGCAAGAGG 354
```

```
DB 372 CAGCTCAGCAGCTGACATCTGAAGACTCTGGGCTCTATTTCTGTGCAAGAGTG 425
|||||
RESULT 13
AAV03926
ID AAV03926 standard; cDNA; 491 BP.
XX
XX AAV03926;
XX
XX 01-JUN-1998 (first entry)
XX
XX Mouse 2H7 antibody heavy chain variable region cDNA.
XX
XX Mouse; murine; heavy chain; variable region;
XX immunoglobulin fragment production; Ig fragment production;
XX monoclonal antibody 2H7; human B-cell surface antigen; ss.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX sig_peptide 72..128
XX /*tag= a
XX mat_peptide 129..491
XX /*tag= b
XX
XX US5693493-A.
XX
XX 02-DEC-1997.
XX
XX 25-MAY-1995; 95US-0450731.
XX
XX 29-MAR-1990; 90US-0501092.
XX 01-NOV-1985; 85US-0793980.
XX 27-OCT-1986; 86WO-US02269.
XX 24-JUL-1987; 87US-0077528.
XX 11-JAN-1988; 88US-0142039.
XX 08-DEC-1992; 92US-0987555.
XX 18-AUG-1994; 94US-0299085.
XX 25-MAY-1995; 95US-0450731.
XX
XX (XOMA ) XOMA CORP.
XX
XX Better M, Horwitz AH, Lei S, Liu AY, Robinson RR;
XX Wall R, Wilcox GL;
XX
XX WPI; 1998-031749/03.
XX P-PSDB; AAW41070.
XX
XX Production of chimeric antibody fragments - by culturing E. coli
XX transformed with dicistronic expression cassette
XX
XX Example IV; Fig 21; 98pp; English.
XX
XX The present sequence was used in the development of a novel method
XX for the production of an immunoglobulin (Ig) fragment capable of
XX binding an antigen. The method comprises culturing an E. coli host
XX that has been transformed with a nucleic acid molecule encoding the
XX Ig fragment, under conditions so that the Ig fragment is produced
XX and secreted. The nucleic acid molecule comprises DNA sequences
XX encoding: (a) pectate lyase secretion signal sequence operably
XX linked to a DNA sequence encoding at least the variable region of
XX an Ig Fd molecule; and (b) pectate lyase secretion signal sequence
XX operably linked to a DNA sequence encoding at least the variable
XX region of an Ig light chain, where (a) and (b) are operably linked
XX to a single prokaryotic promoter to form a dicistronic
XX transcription unit. The method is used to produce chimeric Fab
XX molecules, e.g. derived from murine monoclonal antibody 2H7 raised
XX against human B-cell surface antigen. The invention provides a
XX novel approach for producing genetically engineered antibodies of
XX desired variable region specificity and constant region
XX properties. The cloned Ig gene products can be produced by
XX expression in genetically engineered organisms. The application of
```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 17:29:10 ; Search time 168.759 seconds  
(without alignments)  
6267.222 Million cell updates/sec

Title: US-08-836-455-3

Perfect score: 461

Sequence: 1 ATGGAATGCAGCTGGTCTT.....CTGTCCTCGAAGCTTGGG 461

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA:

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	461	100.0	461 9	US-09-861-294-3
2	461	100.0	461 12	US-10-367-506-3
3	327.2	71.0	409 14	US-10-160-232-7
4	326.6	70.8	736 13	US-10-006-773-12
5	310	67.2	1570 11	US-09-795-515-6
6	297.2	64.5	9199 11	US-09-911-692-3
7	297.2	64.5	9209 11	US-09-911-703-3
8	297.2	64.5	9209 11	US-09-905-928-2
9	297.2	64.5	9209 12	US-10-238-681-3
10	297.2	64.5	9209 14	US-10-096-964-2
11	297.2	64.5	18986 13	US-10-109-853-2
12	296.4	64.0	1392 14	US-10-216-484-8
13	295.2	64.0	482 9	US-09-881-823-19
14	292	63.3	478 13	US-10-040-739-911
15	291.4	63.2	1314 10	US-09-903-327A-5
16	291.4	63.2	1516 10	US-09-903-327A-1
				Sequence 3, Appli
				Sequence 3, Appli
				Sequence 7, Appli
				Sequence 12, Appli
				Sequence 6, Appli
				Sequence 3, Appli
				Sequence 3, Appli
				Sequence 2, Appli
				Sequence 2, Appli
				Sequence 8, Appli
				Sequence 19, Appli
				Sequence 911, Appli
				Sequence 5, Appli

Sequence 77, Appli  
Sequence 85, Appli  
Sequence 85, Appli  
Sequence 87, Appli  
Sequence 9, Appli  
Sequence 5, Appli  
Sequence 10, Appli  
Sequence 5, Appli  
Sequence 9, Appli  
Sequence 234, App  
Sequence 304, App  
Sequence 298, App  
Sequence 285, App  
Sequence 349, App  
Sequence 232, App  
Sequence 233, App  
Sequence 380, App  
Sequence 382, App  
Sequence 383, App  
Sequence 384, App  
Sequence 239, App  
Sequence 273, App  
Sequence 351, App  
Sequence 381, App  
Sequence 393, App  
Sequence 394, App  
Sequence 395, App  
Sequence 11, Appli  
Sequence 11, Appli

US-09-753-436-77  
US-10-244-821-85  
US-10-150-762-85  
US-10-150-762-87  
US-09-911-703-9  
US-09-905-928-5  
US-10-238-681-10  
US-10-096-964-5  
US-09-911-692-9  
US-10-207-655-234  
US-10-207-655-304  
US-10-207-655-298  
US-10-207-655-285  
US-10-207-655-349  
US-10-207-655-232  
US-10-207-655-233  
US-10-207-655-380  
US-10-207-655-382  
US-10-207-655-383  
US-10-207-655-384  
US-10-207-655-239  
US-10-207-655-273  
US-10-207-655-351  
US-10-207-655-381  
US-10-207-655-393  
US-10-207-655-394  
US-10-207-655-395  
US-09-897-006-11  
US-09-897-511A-11

17 289.6 62.8 422 9  
18 289.2 62.7 406 12  
19 289.2 62.7 406 14  
20 289.2 62.7 406 14  
21 283.6 61.5 420 11  
22 283.6 61.5 420 11  
23 283.6 61.5 420 12  
24 283.6 61.5 420 14  
25 283.6 61.5 453 11  
26 281.6 61.1 1536 14  
27 280.2 60.8 1518 14  
28 280.2 60.8 1572 14  
29 280.2 60.8 1584 14  
30 279.8 60.7 2325 14  
31 279.6 60.7 1527 14  
32 279 60.5 1566 14  
33 278.8 60.5 1515 14  
34 278.8 60.5 1515 14  
35 278.8 60.5 1515 14  
36 278.8 60.5 1515 14  
37 278.8 60.5 1521 14  
38 278.8 60.5 1521 14  
39 278.8 60.5 1521 14  
40 278.8 60.5 1521 14  
41 278.8 60.5 1521 14  
42 278.8 60.5 1521 14  
43 278.8 60.5 1521 14  
44 278.6 60.4 5691 10  
45 278.6 60.4 5691 11

## ALIGNMENTS

## RESULT 1

US-09-861-294-3  
; Sequence 3, Application US/09861294  
; Patent No. US20020098190A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; FILE OF INVENTION: TUMORS BEARING HMEF AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/09/861.294  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 461  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(461)  
; NAME/KEY: sig\_peptide  
; LOCATION: (1)...(57)  
; NAME/KEY: mat\_peptide  
; LOCATION: (58)...(461)  
US-09-861-294-3

Query Match 100.0%; Score 461; DB 9; Length 461;  
Best Local Similarity 100.0%; Pred. No. 5.4e-142;  
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGGAATGCAGCTGGTCTTCTCTCTCTGTCATAACTACAGCTGTCCACTCCAG 60

Db 1 ATGGAATGCAGCTGGTCTTCTCTCTCTGTCATAACTACAGCTGTCCACTCCAG 60









COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICANT: Patentin Release #1.0, Version #1.30  
FILING DATE: 25-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/149,099  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-014  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9209 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-911-703-3

Query Match 64.5%; Score 297.2; DB 10; Length 9209;  
Best Local Similarity 79.9%; Pred. No. 2.7e-87;  
Matches 366; Conservative 0; Mismatches 83; Indels 9; Gaps 1;  
QY 1 ATGGAATGACGCTGGCTCTTCTCTCCCTGCTCAATAACTACAGGTGTCCACTCCAG 60  
DB 2401 ATGGGTGGAGCCTCATCTTCTCTCTCTGCTGCTACGGGTGTCTGCTCCAG 2460  
QY 61 GCTTATCTACAGCAGTCTGGGGCTGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 120  
DB 2461 GTACAACTGCAGCAGCCTGGGGCTGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 2520  
QY 121 TGCAGGCTCTGGCTACACATGACAGTACAGTACAGTACAGTACAGTACAGTACAGT 180  
DB 2521 TGCAGGCTCTGGCTACACATGACAGTACAGTACAGTACAGTACAGTACAGTACAGT 2580  
QY 181 GGACAGGCTGGAATGGATGGAATATTTTCTGGAAATGTTCTGGAAATGTTCTGGAAAT 240  
DB 2581 GGTGGGGCTGGAATGGATGGAATATTTTCTGGAAATGTTCTGGAAATGTTCTGGAAAT 2640  
QY 241 CAGAAGTTAAGGGAAGGCTCATTTGACTGCAGACATCTCCAGCAGCAGCCTACATG 300  
DB 2641 CAGAAGTTAAGGGAAGGCTCATTTGACTGCAGACATCTCCAGCAGCAGCCTACATG 2700  
QY 301 CAGATCAGCAGCCTGACATCTGAAGACTCTGGGGTCTATTCTGTGCAAGAGGGAAGTGG 360  
DB 2701 CAGCTCAGCAGCCTGACATCTGAGGACTCTGGGGTCTATTCTGTGCAAGAGTGGTAC 2760  
QY 361 GAGGGTGCT-----CTGACTACTGGGGTCAAGGAACCTCAGTACCCGCTCTCCCTCA 411  
DB 2761 TAGCGGGGTGACTGGTACTTCAATGCTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820  
QY 412 GCCAAAGACGACCCCGCTATCCATCTGGTCC 449  
DB 2821 GCTAGACCAAGGGCCCATCGGTCTTCCCGCTGGCACC 2858

## RESULT 8

US-09-905-928-2

; Sequence 2, Application US/0905928

Publication No. US20030021781A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Darrell R.  
APPLICANT: Hanna, Nabil  
APPLICANT: Leonard, John E.  
APPLICANT: Newman, Roland A.  
APPLICANT: Reff, Mitchell E.  
APPLICANT: Rastetter, William H.  
TITLE OF INVENTION: Therapeutic Application of Chimeric and  
Radiolabeled Antibodies to Human B Lymphocyte Restricted  
Differentiation Antigen for the Treatment of B-Cell Lymphom  
TITLE OF INVENTION: Differentiation Antigen for the Treatment of B-Cell Lymphom  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: 699 Prince St.  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/905,928  
FILING DATE: 17-JUL-2001  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/475,813  
FILING DATE: 07-JUN-1995  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,099  
FILING DATE: 03-NOV-1993  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/978,891  
FILING DATE: 13-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-158  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9209 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: anti-CD20 in TCAE 8  
US-09-905-928-2

Query Match 64.5%; Score 297.2; DB 11; Length 9209;  
Best Local Similarity 79.9%; Pred. No. 2.7e-87;  
Matches 366; Conservative 0; Mismatches 83; Indels 9; Gaps 1;  
QY 1 ATGGAATGACGCTGGGCTCTTCTCTCCCTGCTCAATAACTACAGGTGTCCACTCCAG 60  
DB 2401 ATGGGTGGAGCCTCATCTTCTCTCTGCTGCTTGTCTGCTGCTGCTGCTGCTGCTGCT 2460  
QY 61 GCTTATCTACAGCAGTCTGGGGCTGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 120  
DB 2461 GTACAACTGCAGCAGCCTGGGGCTGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 2520  
QY 121 TGCAGGCTCTGGCTACACATGACAGTACAGTACAGTACAGTACAGTACAGTACAGT 180  
DB 2521 TGCAGGCTCTGGCTACACATGACAGTACAGTACAGTACAGTACAGTACAGTACAGT 2580







```

RESULT 15
US-09-903-327A-5
; Sequence 5, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erguang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGETED
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2003, 16:08:05 ; Search time 1603.21 Seconds  
(without alignments)  
6988.707 Million cell updates/sec

Title: US-08-836-455-3

Perfect score: 461

Sequence: 1 ATGGATGACGTGGTCTT.....CTGGTCCCGAAGCTGGG 461

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_estfun:\*  
15: em\_estom:\*  
16: em\_gss\_hum:\*  
17: em\_gss\_inv:\*  
18: em\_gss\_pin:\*  
19: em\_gss\_vrt:\*  
20: em\_gss\_fun:\*  
21: em\_gss\_nam:\*  
22: em\_gss\_mus:\*  
23: em\_gss\_pro:\*  
24: em\_gss\_rod:\*  
25: em\_gss\_phg:\*  
26: em\_gss\_vrl:\*  
27: gb\_gss1:\*  
28: gb\_gss2:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	337.8	73.3	739	10	BE284158 601099428
2	326.8	70.9	892	10	BG518664 602578528
3	326.6	70.8	637	14	BY724721 BY724721
4	320.8	69.6	941	10	BF138189 601781314

5	320.2	69.5	667	14	BY724790
6	319.6	69.3	1576	11	AK007918 Mus muscu
7	316.6	68.7	767	10	BF165456 601777347
8	316.4	68.6	794	12	BI150371 602915205
9	315.6	68.5	745	12	BG871607 602790090
10	311.2	67.5	546	14	CA576594 K0647806-
11	310.6	67.2	771	10	BF140551 601787584
12	309.6	67.4	529	14	CA576389 K0644A12-
13	308.6	66.9	616	10	BF136295 601781413
14	308.6	66.9	748	10	BF136397 601783227
15	306.6	66.5	700	10	BF540088 602050325
16	305	66.2	644	10	BE281961 601099518
17	305	66.2	793	10	BF136093 601783539
18	303	65.7	891	10	BF138460 601782916
19	303	65.7	955	13	BQ947728 AGENCOURT
20	302.6	65.6	469	14	CA574907 K0622H02-
21	301.6	65.4	690	12	BG968682 602836513
22	301.6	65.4	724	12	BG962137 602826902
23	299.8	65.0	604	10	BF140035 601791175
24	298.4	64.7	538	14	CA578968 K0731A04-
25	297.4	64.5	540	14	CA579750 K0741H07-
26	296.8	64.4	724	10	BF168514 601775412
27	296.8	64.4	847	10	BF165486 601777393
28	296.2	64.3	507	10	BF015548 uy23a08.Y
29	295.8	64.2	519	14	CA580087 K0746G01-
30	295.2	64.0	769	10	BF168856 601775314
31	293.6	63.7	819	12	BG966397 602832896
32	293	63.6	477	14	CA570382 K0505G11-
33	292.8	63.5	640	13	BQ109114 imageqc_6
34	292.6	63.5	485	14	CA578116 K0718G11-
35	292.4	63.4	475	14	CA578534 K0725B11-
36	292.2	63.4	911	10	BE286958 601092470
37	292	63.3	429	14	CA579143 K0733D01-
38	292	63.3	1003	14	CB236612 AGENCOURT
39	291.4	63.2	529	14	CA577954 K0716F08-
40	290.8	63.1	896	13	BUS24012 AGENCOURT
41	290.6	63.0	508	14	CA577080 K0704C12-
42	290.6	63.0	538	14	CA580316 K0749H01-
43	290.4	63.0	598	12	BI104341 602889919
44	290	62.9	1012	10	BF579009 602096127
45	289.6	62.8	672	10	BF580337 602097080

#### ALIGNMENTS

RESULT 1  
BE284158  
LOCUS 601099428F1 NCI\_CGAP\_Lu29 Mus musculus cdna clone IMAGE:3491766 5',  
DEFINITION mRNA sequence.  
ACCESSION BE284158  
VERSION BE284158.1 GI:9160900  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 739)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM8536 row: h column: 07  
High quality sequence stop: 535.

REFERENCE	Mammal J. Sci. Mammalia; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 892)
TITLE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a>
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
LOCUS	BY724721
DEFINITION	musculus cDNA clone A53001123 5', mRNA sequence.
ACCESSION	BY724721
VERSION	BY724721.1



```

KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM

REFERENCE
1 (bases 1 to 941)
TITLE      NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL
COMMENT    National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-f@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM9245 row: k column: 04
High quality sequence stop: 637.
Location/Qualifiers
1. 941
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4009467"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI-CGAP Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; transgenic model WNR-1, expression driven by
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT      232 a 266 c 248 g 195 t
ORIGIN

Query Match      69.6%; Score 320.8; DB 10; Length 941;
Best Local Similarity 84.2%; Pred. No. 1.4e-80;
Matches 377; Conservative 0; Mismatches 62; Indels 9; Gaps 1;

Qy 13 TGGTCTTTCTTCCTCCCTGTCGAGTGTGAGGCTGCGGCGCTCAGTGAAGATGTCCTGCAAGGCTTCT 132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20 TGTATCTTTCTTCCTCCCTGTCGAGTGTGAGGCTGCGGCGCTCAGTGAAGATGTCCTGCAAGGCTTCT 79

Qy 73 CAGTCTGGGCGTGAAGTGTGAGTGTGAGGCTGCGGCGCTCAGTGAAGATGTCCTGCAAGGCTTCT 132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 80 CAGTCTGGGCGTGAAGTGTGAGTGTGAGGCTGCGGCGCTCAGTGAAGATGTCCTGCAAGGCTTCT 139

Qy 133 GCCTACACATGACCACTGTTACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 192
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 140 GGCTACGATCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 199

Qy 193 GAATGGATTGGAATATTTTCTCGGAATGGTGTACTTACTACATACAGAGTTTAAAG 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 GAGTGGATTGACGAGTTATTCCTGGAGTGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 259

Qy 253 GGCAAGGCCCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 GGCAAGGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319

Qy 313 CTGACATCTGAAGACTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 CTGACATCTGAGGACTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 379

Qy 366 --TGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 AATGCTATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 439

Qy 424 CCCCACCCCTCTATCCACTGGTCCCTG 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 440 GCCCATCGTCTCTATCCACTGGTCCCTG 467

```

RESULT 5  
BY724790  
LOCUS  
DEFINITION  
MUSCULUS CDNA clone A530085L22 5', mRNA sequence.  
BY724790  
BY724790.1 GI:27137917  
EST.  
Mus musculus (house mouse)

REFERENCE  
1 (bases 1 to 667)  
AUTHORS

CONTACT: Robert Strausberg, Ph.D.  
Email: cgaabs-f@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov  
Plate: LAM9245 row: k column: 04  
High quality sequence stop: 637.

Location/Qualifiers  
1. 941  
/organism="Mus musculus"

/mol\_type="mRNA"  
/strain="CZECH II"

/db\_xref="taxon:10090"  
/clone="IMAGE:4009467"

/tissue\_type="tumor, metastatic to mammary"  
/lab\_host="DH10B"

/clone\_lib="NCI-CGAP Lu30"  
/notes="Organ: lung; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; transgenic model WNR-1, expression driven by  
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo  
dT. Library constructed by Life Technologies.  
Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 232 a 266 c 248 g 195 t  
ORIGIN

Query Match 69.6%; Score 320.8; DB 10; Length 941;  
Best Local Similarity 84.2%; Pred. No. 1.4e-80;  
Matches 377; Conservative 0; Mismatches 62; Indels 9; Gaps 1;

Qy 13 TGGTCTTTCTTCCTCCCTGTCGAGTGTGAGGCTGCGGCGCTCAGTGAAGATGTCCTGCAAGGCTTCT 132  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 20 TGTATCTTTCTTCCTCCCTGTCGAGTGTGAGGCTGCGGCGCTCAGTGAAGATGTCCTGCAAGGCTTCT 79

Qy 73 CAGTCTGGGCGTGAAGTGTGAGTGTGAGGCTGCGGCGCTCAGTGAAGATGTCCTGCAAGGCTTCT 132  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 80 CAGTCTGGGCGTGAAGTGTGAGTGTGAGGCTGCGGCGCTCAGTGAAGATGTCCTGCAAGGCTTCT 139

Qy 133 GCCTACACATGACCACTGTTACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 192  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 140 GGCTACGATCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 199

Qy 193 GAATGGATTGGAATATTTTCTCGGAATGGTGTACTTACTACATACAGAGTTTAAAG 252  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 200 GAGTGGATTGACGAGTTATTCCTGGAGTGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 259

Qy 253 GGCAAGGCCCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 260 GGCAAGGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319

Qy 313 CTGACATCTGAAGACTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 365  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 320 CTGACATCTGAGGACTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 379

Qy 366 --TGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 380 AATGCTATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 439

Qy 424 CCCCACCCCTCTATCCACTGGTCCCTG 451  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 440 GCCCATCGTCTCTATCCACTGGTCCCTG 467

CONTACT: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/

ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda  
S., Hashizume, W., Hayashida, K., Hirozane, F., Hori, F., Imotani, K.,  
Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno  
H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,  
Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H.,  
Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y.,  
Waki, K., Wachihi, A., Muramatsu, M. and Hayashizaki, Y. Direct  
Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a

mus musculus (house mouse)

mus musculus

mus musculus

mus musculus





/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"  
BASE COUNT 186 a 214 c 190 g 177 t  
ORIGIN

Query Match 68.7%; Score 316.6; DB 10; Length 767;  
Best Local Similarity 84.0%; Pred. No. 2e-79;  
Matches 382; Conservative 0; Mismatches 69; Indels 4; Gaps 2;  
QY 1 ATGGATGACGCTGGGCTTTCTTCTCCCTCTCAATACTACAGGTGCCACTCCAG 60  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
35 ATGGAATGGAGCTGCATCTTCTTCTCCCTCTCAGTAACAGCAGGTGCCACTCGAG 94  
QY 61 GCTTATCTACA-CGAGTCTGGGCTGAGCTGGTGGCTCGAGTGGGCTCGAGTGGTGC 119  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
95 GTCCAGCTTCAGCAGCTCTGGAGCTGAGCTGGTGGGCTGGTCTCAGTGAAGCTGTC 154  
QY 120 CTGCAAGGCTTCTGGGTACACATTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 179  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
155 CTGCAAGACTTCTGGATATACATTCACAGCTATTATATAAAGTGGTGAAGCAGAGCC 214  
QY 180 TGCACAGGCTGGAATGGATGGAAATATTTTCTGGAAATGGTGGTGGTGGTGGTGG 239  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
215 TGCACAGGCTGGAATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGG 274  
QY 240 TCAGAAAGTTTAAAGGCAAGGCTTCATTGAGCTGCAGACACATCTCCAGCAGCCTACAT 299  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
275 TGAGAAAGTTCAAGGCAAGGCTTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 334  
QY 300 GCAGATCAGCAGCTGACATCTGAAGTCTGCGGTCTATTTCTGTCGAAGA---GGGAA 356  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
335 GCAGCTCAGAAGCTGACATCTCAGAAGCTGCGGTCTATTTCTGTCGAAGAAGCGGTGG 394  
QY 357 CTGGAGGCTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 416  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
395 TTACGAGGTTTACTTTGACTACTTGGGCAAGGCAAGCAGTATCAGTCTGCTCAGGCCAA 454  
QY 417 AACGACACCCCAACCCGCTTATCCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 451  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
455 AACACAGCCCAACCGCTCTATCCACTGGCCCTG 489

RESULT 8  
BI150371  
LOCUS 602915205F1 NCI\_CGAP\_Lu29 Mus musculus cDNA clone IMAGE:5065801 5',  
DEFINITION mRNA sequence.  
ACCESSION BI150371 GI:14610372  
VERSION BI150371.1  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 794)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L1A11177 row: 1 column: 02  
High quality sequence stop: 794.  
Location/Qualifiers  
1. 794

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="czech II"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5065801"  
/tissue\_type="spontaneous tumor, metastatic to mammary."  
Stem cell origin.  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Lu29"  
/notes="Organ: lung; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"  
BASE COUNT 185 a 229 c 200 g 180 t  
ORIGIN

Query Match 68.6%; Score 316.4; DB 12; Length 794;  
Best Local Similarity 82.6%; Pred. No. 2.3e-79;  
Matches 375; Conservative 0; Mismatches 76; Indels 3; Gaps 1;  
QY 1 ATGGAATGACGCTGGGCTTTCTTCTCTCTCTCAATACTACAGGTGCCACTCCAG 60  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
42 ATGGATGAGCTGGATCTTCTTCTCTCTCTCTCAGAACTGAGGTGCCATGCGAG 101  
QY 61 GCTTATCTACAGCAGCTGGGCTGAGCTGGTGGGCTCAGTGGGCTCAGTGAAGATGCC 120  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
102 GTCAGCTGCAGCAGCTGGACCTGAGCTGGTGAAGCTGGGCTTCAGTGAAGATATCC 161  
QY 121 TGCAGGCTTCTGGCTACACATTCACCAGTTACAAATGACATGAGTGGTAAAGCAGACCT 180  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
162 TGCAGGCTTCTGGCTACAGCTTCACAAGCTATATATACATGAGTGGTGAAGAGGCTCT 221  
QY 181 GGACAGGCTTGAAGTGGATTTGAAATATATTTTCTGGAATGCTGATCTTACTACAAT 240  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
222 GGACAGGCTTGAAGTGGATTTGAAATATATTTTCTGGAATGCTGATCTTACTACAAT 281  
QY 241 CAGAAAGTTAAGGCAAGGCTCATTTGACGTGACAGACATCTCCAGCAGCAGCTCATG 300  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
282 GAGAAAGTTAAGGCAAGGCTCATTTGACGTGACAGACATCTCCAGCAGCAGCTCATG 341  
QY 301 CAGATCAGCAGCTGACATCTGAGACTCTGCGGTCTATTCTGTCGAAGAGGAACT-- 358  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
342 CAGCTCAGCAGCTGACATCTGAGACTCTGAGACTCTGAGACTCTGAGACTCTGAGACTCT 401  
QY 359 -GGGAGGCTCTGAGCTACTGGGCTCAAGGAACTCTGAGCTCAGCTCAGCTCCTCAGCCAAA 417  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
402 AGTTACGACGGTTTCTTACTGGGCCAAGGACTCTGCTGCTACTGCTCTGCGAGCCAAA 461  
QY 418 ACAGACACCCCAACCCGCTCTATCCACTGGTCCCTG 451  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
462 ACAACAGCCCAACCGTCTATCCACTGGCCCTG 495

RESULT 9  
BG871607  
LOCUS 602790090F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4921295 5',  
DEFINITION mRNA sequence.  
ACCESSION BG871607 GI:14222147  
VERSION BG871607.1  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 745)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.













```

||||| ||||||| ||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 271 GAGAGTTCAAGGCAAGGCCACATTCACTGCAGATACATCTCCACACAGCCTACATG 330
Qy 301 CAGATCAGCAGCCTGACATCTGAAGACTCTGCGTCTATTCTGTGCAAGAGGGAAGTGG 360
Db 331 CAACTCAGCAGCCTGACATCTGAGGACTCTGCCGTCTATTACTGTGCAAGACGGCTGGGA 390
Qy 361 GAGGTGCTCTGGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCCAAAACG 420
Db 391 CGATGGTACTTCGATGTCTGGGGCGCAGGACCCACGGTCACCGTCTCTCAGAGATCAG 450
Qy 421 ACACCCCCACCGTCTATCCACTGGTCCC 449
Db 451 TCCTTCCCAAATGTCTTTCCCGCTGCTC 479
```

Search completed: August 30, 2003, 19:44:51  
Job time : 1608.21 secs



```

FT Region /note= "complementarity determining region 2"
FT 86..117
FT /label= FR3
FT /note= "framework region 3"
FT 118..126
FT /label= CDR3
FT /note= "complementarity determining region 3"
FT 127..137
FT /label= FR1
FT /note= "framework region 4"
FT 138..153
FT /label= Constant
XX
XX W09722699-A2.
XX
XX 26-JUN-1997.
XX
XX 19-DEC-1996; 96WO-US20757.
XX
XX 20-DEC-1995; 95US-0575762.
XX 26-JAN-1996; 96US-0591965.
XX 13-DEC-1996; 96US-0766350.
XX
XX (KENT ) UNIV KENTUCKY.
XX
XX Chatterjee M, Chatterjee SK, Foon KA;
XX WPI; 1997-341690/31.
XX N-PSDB; AAT85150.
XX
XX Monoclonal anti-idiotypic antibody 11D10 - elicits immune response
XX against human milk fat globule disease associated tumours,
XX especially breast cancer
XX
XX Claim 10; Page 94-95; 130pp; English.
XX
XX This polypeptide sequence comprises the heavy chain variable region
XX (VH) of monoclonal anti-idiotypic antibody 11D10 produced by
XX hybridoma cell line ATCC 12020. 11D10 was obtained by immunising
XX naive mice with MC-10 anti-HMFG antibody to obtain an anti-idiotypic
XX response. It elicits an immune response against a specific epitope
XX of a high mol.wt. mucin of human milk fat globule (HMFG). It
XX induces an immunological response to HMFG in mice, rabbits, monkeys
XX and patients with advanced HMFG-associated tumours. Pharmaceutical
XX compositions and vaccines comprising 11D10, 11D10 polypeptides
XX and/or 11D10 polynucleotides (see also AAT85149-50) are claimed.
XX Also claimed are diagnostic kits and methods of using 11D10, 11D10
XX polypeptides and/or 11D10 polynucleotides, including methods of
XX treating HMFG-associated tumours. 11D10 is also used in a claimed
XX method of palliating HMFG-associated disease and in claimed kits to
XX detect or quantify anti-HMFG antibody.
XX (Updated on 25-MAR-2003 to correct PR field.)
XX
XX Sequence 153 AA;
XX
XX Query Match 100.0%; Score 153; DB 18; Length 153;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-130;
XX Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MECSWFLFLSLTTGVHSGAYLQSGAEIVRSGASVKMSKASGYTLTSTNNHWVKQTP 60
DB 1 MECSWFLFLSLTTGVHSGAYLQSGAEIVRSGASVKMSKASGYTLTSTNNHWVKQTP 60
QY 61 GQGLEWIGNIFPGNGDTYYNQRKFKASLTADTSSSTAYMQISLTSSESAVYFCARGNW 120
DB 61 GQGLEWIGNIFPGNGDTYYNQRKFKASLTADTSSSTAYMQISLTSSESAVYFCARGNW 120
QY 121 EGALDWGQGTSTVSSAKTTPPPVPLVPGSL 153
DB 121 EGALDWGQGTSTVSSAKTTPPPVPLVPGSL 153

```

RESULT 2

```

AAW87594
ID AAW87594 standard; Protein; 153 AA.
XX
XX AAW87594;
XX
XX 16-MAR-1999 (first entry)
XX
XX Antibody 11D10 heavy chain variable region.
XX
XX Murine; mouse; antibody; heavy chain; variable region; anti-idiotypic;
XX human milk fat globule; tumour; ovary; lung; pancreas; carcinoma; breast.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX Peptide 1..19
XX /note= "signal peptide"
XX Protein 20..153
XX /note= "mature protein"
XX Region 20..49
XX /label= framework_1
XX Domain 50..54
XX /label= CDR1
XX /note= "complementarity determining region 1"
XX Region 55..68
XX /label= framework_2
XX Domain 69..85
XX /label= CDR2
XX /note= "complementarity determining region 2"
XX Region 86..117
XX /label= framework_3
XX Domain 118..126
XX /label= CDR3
XX /note= "complementarity determining region 3"
XX Region 127..137
XX /label= framework_4
XX
XX W09856419-A1.
XX
XX 17-DEC-1998.
XX
XX 12-JUN-1998; 98WO-US12250.
XX
XX 11-JUN-1998; 98US-0096244.
XX 13-JUN-1997; 97US-0049540.
XX
XX (KENT ) UNIV KENTUCKY RES FOUND.
XX
XX Chatterjee M, Foon KA;
XX WPI; 1999-060029/05.
XX N-PSDB; AAV83773.
XX
XX Delaying development of, or treating, HMFG-associated tumours -
XX using anti-idiotypic antibody 11D10 raised against antibodies to
XX human milk fat globule protein
XX
XX Disclosure; Fig 2; 54pp; English.
XX
XX This sequence represents the heavy chain variable region of the murine
XX antibody 11D10. This anti-idiotypic antibody is used to delay the
XX development of, or treat, a human milk fat globule (HMFG) associated
XX tumour in an individual having low tumour burden. The antibody 11D10
XX is used to prevent the recurrence of HMFG-associated tumours e.g.
XX ovarian, non-small cell lung and pancreatic carcinoma, especially for
XX treating breast tumours.
XX
XX Sequence 153 AA;
XX
XX Query Match 100.0%; Score 153; DB 20; Length 153;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-130;
XX Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MECSWFLFLLSITTCVHSQAYLQSGAELVRSGASVKMSCKASGYTLTSYNMHWKQTP 60  
 Db 1 MECSWFLFLLSITTCVHSQAYLQSGAELVRSGASVKMSCKASGYTLTSYNMHWKQTP 60  
 QY 61 GQGLEWIGNIFPGNGDTYYNQKFKGASLTADTSSSTAYMQISSLTSEDSAVYFCARGNW 120  
 Db 61 GQGLEWIGNIFPGNGDTYYNQKFKGASLTADTSSSTAYMQISSLTSEDSAVYFCARGNW 120  
 QY 121 EGALDYWGQGTSTVTSSAKTTTPPPVPLVPGSL 153  
 Db 121 EGALDYWGQGTSTVTSSAKTTTPPPVPLVPGSL 153

RESULT 3  
 AAO16293  
 ID AAO16293 standard; Protein; 153 AA.  
 XX  
 AC AAO16293;  
 DT 20-MAR-2003 (first entry)  
 XX  
 DE Mouse 11D10 antibody heavy chain variable region.  
 XX  
 KW Mouse; murine; vaccine; tumour; human milk fat globules; HMFG;  
 KW carcinoembryonic antigen; CEA; 11D10; 3H1; HMFG-associated tumour;  
 KW CEA-associated tumour; anti-idiotypic antibody.  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..19  
 FT /label= signal\_peptide  
 FT Protein 20..153  
 FT /note= "Mature murine 11D10 antibody heavy chain variable  
 region"  
 XX  
 PN WO200292012-A2.  
 XX  
 PD 21-NOV-2002.  
 XX  
 PF 17-MAY-2002; 2002WO-US15840.  
 XX  
 PR 17-MAY-2001; 2001US-0861294.  
 XX  
 PA (KENT ) UNIV KENTUCKY RES FOUND.  
 XX  
 PI Chatterjee M, Foon KA;  
 DR WPI; 2003-129216/12.  
 DR N-PSDB; AAL51274.  
 XX  
 PT Use of anti-idiotypic antibodies for human milk fat globules (HMFG) - or  
 PT carcinoembryonic antigen (CEA)-associated tumor for delaying the  
 PT development of, or treating a HMFG- or CEA-associated tumor (e.g.  
 PT breast tumor) in humans  
 XX  
 PS Claim 2; Fig 2; 98pp; English.  
 XX  
 CC The invention comprises a method for delaying the development of, or  
 CC treating a tumour that is associated with human milk fat globules (HMFG)  
 CC or carcinoembryonic antigen (CEA). The method of the invention involves  
 CC administering an anti-idiotypic antibody for HMFG (e.g. 11D10), and an  
 CC anti-idiotypic antibody for CEA (e.g. 3H1). The method is useful for  
 CC delaying the development, of or treating HMFG/CEA-associated tumours. The  
 CC present amino acid sequence represents the heavy chain variable region of  
 CC the mouse 11D10 anti-idiotypic antibody.  
 XX  
 SQ Sequence 153 AA;  
 Query Match 100.0%; Score 153; DB 24; Length 153;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-130;  
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECSWFLFLLSITTCVHSQAYLQSGAELVRSGASVKMSCKASGYTLTSYNMHWKQTP 60  
 Db 1 MECSWFLFLLSITTCVHSQAYLQSGAELVRSGASVKMSCKASGYTLTSYNMHWKQTP 60  
 QY 61 GQGLEWIGNIFPGNGDTYYNQKFKGASLTADTSSSTAYMQISSLTSEDSAVYFCARGNW 120  
 Db 61 GQGLEWIGNIFPGNGDTYYNQKFKGASLTADTSSSTAYMQISSLTSEDSAVYFCARGNW 120  
 QY 121 EGALDYWGQGTSTVTSSAKTTTPPPVPLVPGSL 153  
 Db 121 EGALDYWGQGTSTVTSSAKTTTPPPVPLVPGSL 153

RESULT 4  
 AAY49210  
 ID AAY49210 standard; Protein; 152 AA.  
 XX  
 AC AAY49210;  
 DT 07-FEB-2000 (first entry)  
 XX  
 DE MAb 1A7 heavy chain variable region.  
 XX  
 KW Monoclonal antibody; MAb; 1A7; GD2; immune response; melanoma;  
 KW neuroblastoma; glioma; soft tissue carcinoma; small cell carcinoma;  
 KW tumor-associated antigen.  
 XX  
 OS Synthetic.  
 OS Mas sp.  
 XX  
 PN US5977316-A.  
 PD 02-NOV-1999.  
 XX  
 PE 16-JAN-1996; 96US-0591196.  
 XX  
 PR 17-JAN-1995; 95US-0372676.  
 XX  
 PA (KENT ) UNIV KENTUCKY.  
 XX  
 PI Foon KA, Chatterjee SK, Chatterjee M;  
 DR WPI; 1999-619711/53.  
 DR N-PSDB; AA231366.  
 XX  
 PT Monoclonal antibody 1A7 which elicits an anti-GD2 immunological  
 PT response, useful for the development of products for the detection and  
 PT treatment of cancers  
 XX  
 PS Claim 1; Fig 2; 74pp; English.  
 XX  
 CC The invention provides a monoclonal antibody (MAb) designated 1A7, which  
 CC elicits an anti-GD2 (tumor-associated antigen) immunological response in  
 CC humans. MAb 1A7 has defined light and heavy chain variable region  
 CC sequences. The MAb 1A7 and polypeptides can be used for eliciting an  
 CC anti-GD2 immune response. The polypeptides can also be used for detecting  
 CC or purifying anti-GD2 antibody. The products can be used for treating GD2  
 CC -associated diseases, e.g. melanoma, neuroblastoma, glioma, soft tissue  
 CC carcinoma, and small cell carcinoma. They can be used for palliating the  
 CC disease or for reducing the risk of recurrence. The present sequence  
 CC represents the heavy chain variable region of MAb 1A7.  
 XX  
 SQ Sequence 152 AA;  
 Query Match 20.3%; Score 31; DB 20; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-20;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 ALDYWGQGTSTVTSSAKTTTPPPVPLVPGSL 153  
 Db 122 ALDYWGQGTSTVTSSAKTTTPPPVPLVPGSL 152

## RESULT 5

AAY28469 ID AAY28469 standard; Protein; 152 AA.

XX AC AAY28469;

XX DT 06-OCT-1999 (first entry)

XX DE Heavy chain variable region of MAb 1A7.

XX KW Heavy chain variable region; antibody 1A7; T cell response; carcinoma; ganglioside GD2; CDR; complementarity determining region; melanoma.

XX OS Mus musculus.

XX FH Key Location/Qualifiers

XX FT Peptide 1..19

XX FT /label= "Signal peptide"

XX FT Protein 20..152

XX FT /label= "Mature peptide"

XX FT Region 20..49

XX FT /label= "Framework 1"

XX FT Domain 50..54

XX FT /label= "CDR 1"

XX FT /note= "Complementarity determining region 1"

XX FT Region 55..65

XX FT /label= "Framework 2"

XX FT Domain 69..84

XX FT /label= "CDR 2"

XX FT /note= "Complementarity determining region 2"

XX FT Region 85..116

XX FT /label= "Framework 3"

XX FT Domain 117..125

XX FT /label= "CDR 3"

XX FT /note= "Complementarity determining region 3"

XX FT Region 126..136

XX FT /label= "Framework 4"

XX FT Region 137..152

XX FT /label= "Constant region"

XX FT US5935821-A.

XX PN 10-AUG-1999.

XX PD 21-NOV-1996; 96US-0752844.

XX PF 21-NOV-1996; 96US-0752844.

XX PR 17-JAN-1995; 95US-0372676.

XX PR 16-JAN-1996; 96US-0591196.

XX XX (KENT ) UNIV KENTUCKY.

XX PA Chatterjee M, Chatterjee SK, Foon KA;

XX PI WPI; 1999-457600/38.

XX DR N-PSDB; AAX89553.

XX DR

XX XX Anti-GD2 immunological peptides useful for the treatment of tumours especially melanomas and small cell carcinomas

XX PS Claim 1; Fig 2; 84pp; English.

XX CC

This polypeptide is the variable heavy chain region of monoclonal anti-idiotypic antibody 1A7. The polypeptide has three CDRs (complementarity determining regions) and four frame work regions. When administered to an individual the 1A7 antibody induces an immune response against ganglioside GD2. The light chain variable region of the 1A7 antibody (AAY28468) is also capable of eliciting an anti GD2 response in mammals. Both the heavy and light chain variable regions of the 1A7 antibody produce anti-GD2 T cell and antibody responses. The peptides and antibodies may be useful for the modulation of ganglioside GD2, and particularly for the treatment of GD2-associated tumours e.g. melanoma, neuroblastoma, glioma, soft tissue sarcoma, and small cell carcinoma

CC (including small cell lung cancer).

XX SQ Sequence 152 AA;

XX

Query Match 20.3%; Score 31; DB 20; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-20;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 ALDYGQGTSTVTSSAKTTPPPVPLVPGSL 153  
 |||||  
 Db 122 ALDYGQGTSTVTSSAKTTPPPVPLVPGSL 152  
 |||||

RESULT 6  
 AAY21546 ID AAY21546 standard; Protein; 152 AA.

XX AC AAY21546;

XX DT 03-AUG-1999 (first entry)

XX DE Monoclonal antibody 1A7 heavy chain variable region.

XX KW Psoriasis; immunological response; anti-idiotypic antibody; glutamate;  
 chronic plaque; pustular; plaque-type psoriasis; psoriatic arthritis;  
 monoclinal antibody; 1A7.  
 XX KW Unidentified.  
 XX OS WO9925380-A2.  
 XX PN 27-MAY-1999.  
 XX PD 17-NOV-1998; 98WO-US24607.  
 XX PF 16-NOV-1998; 98US-0192838.  
 XX PR 17-NOV-1997; 97US-0065774.  
 XX XX (KENT ) UNIV KENTUCKY RES FOUND.

XX XX Chatterjee M, Foon KA;

XX PI WPI; 1999-347407/29.

XX DR N-PSDB; AAX60630.

XX XX Treatment of psoriasis

XX PS Disclosure; Fig 3; 48pp; English.

XX XX

The invention provides a method of treating of psoriasis by administering an antigen which has similar immunogenic properties to an antigen expressed on cells of psoriatic tissue so that an immunological response is elicited in the individual. The antigen stimulates the generation of anti-idiotypic antibodies that neutralize the aberrant immune response causing the psoriasis. The method is used to treat psoriasis, especially chronic plaque, glutamate, pustular, plaque-type psoriasis or psoriatic arthritis. The compositions allow the individual's own immune system to act against psoriatic tissue. The present sequence represents the heavy chain variable region of monoclonal antibody 1A7.

XX SQ Sequence 152 AA;

XX

Query Match 20.3%; Score 31; DB 20; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-20;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 ALDYGQGTSTVTSSAKTTPPPVPLVPGSL 153  
 |||||  
 Db 122 ALDYGQGTSTVTSSAKTTPPPVPLVPGSL 152  
 |||||

RESULT 7  
 AAY21816



ID AAY21816 standard; Protein: 122 AA.  
XX AC AAY21816;  
XX DT 10-SEP-1999 (first entry)  
XX DE Anti-STX1 heavy chain variable region.  
XX KW Humanised; monoclonal antibody; MAb; Shiga toxin; Immunoglobulin;  
KW Ig constant region; enterohemorrhagic Escherichia coli; EHEC; uremia;  
KW edema; bloody diarrhoea; hemorrhagic colitis; hemolytic uremic syndrome;  
KW thrombocytopenia; EHEC-mediated disease; anti-STX1.  
XX OS Shigella dysenteriae.  
XX FH Key Location/Qualifiers  
XX DE Misc-difference 23 /note= "encoded by AAG"  
XX FT Region 31..35 /note= "complementarity determining region (CDR) 1;  
FT specifically claimed for in claim 10"  
XX FT Region 50..66 /note= "complementarity determining region (CDR) 2;  
FT specifically claimed for in claim 10"  
XX FT Region 99..111 /note= "complementarity determining region (CDR) 3;  
FT specifically claimed for in claim 10"  
XX PN WO9932645-A1.  
XX PD 01-JUL-1999.  
XX XX 22-DEC-1998; 98WO-US27267.  
XX PR 18-DEC-1998; 98US-0215163.  
XX PR 23-DEC-1997; 97US-0068635.  
XX PA (MELT/) MELTON-CELSA A.  
PA (OBRI/) O'BRIEN A D.  
PA (SCHM/) SCHMITT C K.  
PA (STIN/) STINSON J L.  
PA (WONG/) WONG H.  
XX XX Melton-Celsa A, O'Brien AD, Schmitt CK, Stinson JL;  
PI Wong H;  
XX XX WPI: 1999-418935/35.  
DR N-PSDB; AAX82028.  
XX XX Humanized monoclonal antibodies against Shiga toxins, useful for  
PT protection against enterohemorrhagic Escherichia coli or other Shiga  
PT toxin producing bacteria  
XX XX Claim 5; Fig 3; 75pp; English.  
XX XX The invention relates to humanised monoclonal antibodies (MAb) against  
CC Shiga toxins. The humanised MAb that binds to Shiga toxin comprises a  
CC constant and a variable region, where: (a) the constant region contains  
CC at least part of a human immunoglobulin (Ig) constant region; and (b) the  
CC variable region contains at least part of a non-human Ig variable region.  
CC Host cells transformed with vectors encoding a humanised MAb against  
CC Shiga toxin type 2 is useful for treating a patient with an infection  
CC caused by enterohemorrhagic Escherichia coli (EHEC) or other Shiga toxin  
CC producing bacteria. The humanised MAb can also be used to reduce illness  
CC caused by EHEC or other Shiga toxin producing bacteria. EHEC are  
CC associated with food-borne outbreaks of bloody diarrhoea (hemorrhagic  
CC colitis) and the hemolytic uremic syndrome. In particular, the humanised  
CC MAb ameliorate edema, thrombocytopenia and uremia associated with EHEC-  
CC mediated disease. The present sequence represents an anti-STX1 heavy  
CC chain variable region.  
XX XX Sequence 122 AA;

Query Match 17.6%; Score 27; DB 20; Length 122;  
Best Local Similarity 100.0%; Pred. No. 1.3e-16;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 89 LPTADTSSSTAYMQISLTSEDSAVYFC 115  
DB 70 LTADTSSSTAYMQISLTSEDSAVYFC 96  
RESULT 8  
AAW03200  
ID AAW03200 standard; Protein: 152 AA.  
XX AC AAW03200;  
XX XX 25-MAR-2003 (updated)  
DT 26-FEB-1997 (first entry)  
XX XX Anti-idiotypic monoclonal antibody 1A7 variable heavy chain.  
XX DE Murine; mouse; anti-idiotypic; monoclonal antibody; MAB; 1A7;  
XX KW variable heavy chain; ganglioside 2; GD2; 14G2a; neuroblastoma;  
KW glycosphingolipid; human; neuroectodermal; tumour; glioma; lung;  
KW malignant melanoma; soft tissue sarcoma; small cell carcinoma;  
KW vaccine; treatment; palliate; detection; diagnosis;  
KW recombinant production; purification; probe; primer; assay;  
KW amplification; gene therapy.  
XX XX Mus musculus.  
XX FH Key Location/Qualifiers  
XX DE Peptide 1..19  
XX FT Region 20..49 /label= sig\_peptide  
XX FT Region 50..54 /note= "framework region 1"  
XX FT Region 55..68 /note= "complementarity determining region 1"  
XX FT Region 69..84 /note= "framework region 2"  
XX FT Region 85..116 /note= "complementarity determining region 2"  
XX FT Region 117..125 /note= "framework region 3"  
XX FT Region 126..136 /note= "complementarity determining region 3"  
XX FT Misc-difference 125 /note= "framework region 4"  
XX FT /note= "corresponding codon TAC"  
XX XX WO9622373-A2.  
XX XX 25-JUL-1996.  
XX XX 17-JAN-1996; 96WO-US00882.  
XX XX 17-JAN-1995; 95US-0372676.  
XX PR 16-JAN-1996; 96US-0591196.  
XX PR 17-JAN-1995; 95US-0372676.  
XX PR 16-JAN-1996; 96US-0591196.  
XX XX (KENT ) UNIV KENTUCKY.  
XX XX Chatterjee M, Chatterjee SK, Foon KA;  
XX WPI: 1996-354530/35.  
DR N-PSDB; AAT31333.  
XX XX Monoclonal antibody 1A7 and related polynucleotide(s) and  
PT polypeptide(s) - useful to treat or palliate a GD2-associated  
PT disease, e.g. melanoma and glioma  
XX XX Claim 9; Fig 2; 141pp; English.

XX The present sequence is that of the murine anti-idiotypic monoclonal  
 CC antibody (MAB) 1A7 variable heavy chain. MAB 1A7 was raised against  
 CC the anti-ganglioside 2 (GD2) MAB 14G2a, which binds an unique  
 CC epitope of GD2. As the glycosphingolipid GD2 is expressed at high  
 CC density by human neuroectodermal tumours, e.g. malignant melanoma,  
 CC neuroblastoma, glioma, soft tissue sarcoma and small cell carcinoma  
 CC of the lung, MAB 1A7, or its cDNA can be used in a vaccine to treat  
 CC or palliate such diseases. They can also be used to reduce the  
 CC risk of recurrence of a clinically detectable tumour, and detect an  
 CC anti-GD2 Ab bound to a tumour cell.  
 CC MAB 1A7 overcomes immune tolerance and induces an immune response  
 CC against GD2, which comprises anti-GD2 Ab (humoral response) and  
 CC GD2-specific cells (cellular response). It can be used to purify  
 CC anti-1A7 (AB3), anti-GD2 (Ab1) or 14G2a (Ab1), detect anti-1A7 or  
 CC anti-GD2 in a sample or measure the level of cellular anti-1A7 or  
 CC anti-GD2 activity.  
 CC The cDNA can be used in expression systems for 1A7 prodn., and in  
 CC the prepn. of probes and primers to respectively assay for 1A7  
 CC cDNA, and amplify desired polynucleotides for use in gene therapy.  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 XX

SQ Sequence 152 AA;  
 Query Match 17.68; Score 27; DB 17; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-16;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 WGQGTSTVSSAKTTPPVPLVPGSL 153  
 |||||  
 Db 126 WGQGTSTVSSAKTTPPVPLVPGSL 152

## RESULT 9

AAG80224  
 ID AAG80224 standard; protein; 279 AA.

XX AC AAG80224;

XX DT 22-JAN-2002 (first entry)

XX DE Gamma-glutamyl transferase antibody 138H11 antigen binding region.

XX KW Antigen binding region; antibody; 138H11; VH region; VL region;  
 KW human; gamma-glutamyl transferase; gamma-glutamyl transpeptidase;  
 KW EC 2.3.2.2; kidney cell; carcinoma; renal cell; cytostatic; tumour;  
 KW antirheumatic; antiarthritic; liver; stomach; ovary; breast; glioma;  
 KW melanoma; autoimmune disease; rheumatoid arthritis; metabolic disorder;  
 KW leukotriene biosynthesis; mercaptoacid; glutathione; metastasis;  
 KW differentiation; adenocarcinoma.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Domain 1..141

FT Domain 142..279

FT /label= VL

XX DE10020034-A1.

XX PD 31-OCT-2001.

XX PF 22-APR-2000; 2000DE-1020034.

XX PR 22-APR-2000; 2000DE-1020034.

XX PA (FISCHER) FISCHER P.

XX PA (DUEBEL) DUEBEL S.

XX PA (SCHEE) SCHERBERICH J E.

XX PI Duebel S, Schmiedl A;

XX

DR WPI; 2002-000467/01.

XX Agent for diagnosis and treatment of carcinoma, especially renal cell  
 PT carcinoma, comprises binding partner for gamma-glutamyl transferase -

PS Claim 5; Fig 2; 12pp; German.

XX This invention describes a novel agent (A) for diagnosis and treatment of  
 CC carcinoma based on substances (I) that are binding partners for  
 CC gamma-glutamyltransferase (II), or its isoforms or enzymatically active  
 CC variants. The products of the invention have cytostatic, antirheumatic  
 CC and antiarthritic activity. (A) are useful for diagnosis and treatment of  
 CC tumours that express (II), especially renal cell carcinoma but also  
 CC carcinoma of liver, stomach, ovary and breast, glioma and melanoma,  
 CC autoimmune diseases, e.g. rheumatoid arthritis, and metabolic disorders,  
 CC e.g. disturbances of leukotriene biosynthesis, formation of mercaptoacids  
 CC or of glutathione. Diagnostically, (A) are particularly used for  
 CC (differential) diagnosis of tumours (especially by immunoscintigraphy,  
 CC immunohistochemistry or immunoassays, or for detecting multiple drug  
 CC resistance, metastases and relapse, for differentiation of (II)-positive  
 CC tumours (kidney tumours, gliomas or adenocarcinoma of unknown origin).  
 CC Recombinantly produced (A) reduce the human anti-mouse antibody response,  
 CC allowing repeated administration, and can be prepared efficiently as a  
 CC fusion protein in microorganisms, transgenic animals or plants. Also,  
 CC where used for radio-diagnosis or -therapy, recombinant antibodies are  
 CC quickly cleared from the blood and penetrate tumours better. Where (I) is  
 CC coupled to an effector, e.g. interleukin-2, a high local concentration is  
 CC achieved and systemic toxicity reduced. This sequence represents the  
 CC antigen binding region (VH and VL) from antibody 138H11 against  
 CC gamma-glutamyl transferase (also called gamma-glutamyl transpeptidase,  
 CC EC number 2.3.2.2).

XX SQ Sequence 279 AA;

Query Match 13.7%; Score 21; DB 23; Length 279;

Best Local Similarity 100.0%; Pred. No. 7.1e-11;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 ALDYWGQGTSTVSSAKTTPP 143

|||||

Db 109 ALDYWGQGTSTVSSAKTTPP 129

## RESULT 10

AAB30694

ID AAB30694 standard; protein; 412 AA.

XX AC AAB30694;

XX DT 02-APR-2001 (first entry)

XX DE A fusion of anti-CD20 single chain antibody/streptavidin.

XX KW Streptavidin; tumour cell; cancer; adenocarcinoma;

XX KW hematological malignancy; B9E9.

XX OS Synthetic.

OS Streptomyces avidinii.

OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Protein 1..108

/note= "VL chain"

FT Peptide 109..126

/note= "linker"

FT Protein 127..248

/note= "VH chain"

FT Peptide 249..253

/note= "linker"

FT Protein 254..412

/note= "streptavidin"

XX W0200075333-A1.

PN

```

XX PD 14-DEC-2000.
XX PF
XX PP 05-JUN-2000; 2000WO-US15595.
XX PR 07-JUN-1999; 99US-0137900.
XX PR 03-DEC-1999; 99US-0168976.
XX PA (NEOR-) NEORX CORP.
XX PI Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;
XX PD WPI: 2001-091213/10.
XX DR N-PSDB; AAC86563.
XX
XX New vector constructs for expressing genomic streptavidin fusion
XX PT proteins which are useful for targeting tumour cells associated with
XX PT cancer, e.g. adenocarcinomas -
XX PA (NEOR-) NEORX CORP.
XX PI Example 2; Fig 11B; 100pp; English.
XX
XX The present sequence represents a fusion of an anti-CD20 single chain
XX antibody (B9E9) streptavidin. The fusion protein is expressed using
XX vectors of the invention. The specification describes vector constructs
XX for expressing streptavidin fusion proteins. The vector comprises a
XX nucleic acid encoding streptavidin or its functional variant operatively
XX linked to a promoter, and a cloning site for insertion of a second
XX nucleic acid sequence encoding a polypeptide to be fused with
XX streptavidin, interposed between the promoter and the first nucleic
XX acid sequence. Alternatively, the vector construct comprises a nucleic
XX acid, operatively linked to a promoter, encoding a polypeptide to be
XX fused with streptavidin, and a cloning site for insertion of a second
XX nucleic acid encoding at least 129 amino acids of streptavidin or its
XX functional variant. The fusion proteins are useful for targeting tumour
XX cells, particularly tumour cells associated with cancer,
XX e.g. adenocarcinomas or hematological malignancies. The vector construct
XX is useful for expressing of streptavidin fusion proteins. In particular,
XX these are useful as tools for medical diagnostics and therapeutic
XX purposes, e.g. for detecting the presence or absence of, or treating, a
XX target site within a mammalian host.
XX
XX Sequence 412 AA;
XX
XX Query Match 13.1%; Score 20; DB 22; Length 412;
XX Best Local Similarity 100.0%; Pred. No. 7.9e-10;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 49 TSYNHHWVKQTQGGLEWIG 68
XX 157 TSYNHHWVKQTQGGLEWIG 176
XX
XX Db
XX
XX RESULT 11
XX AAB30695
XX ID AAB30695 standard; Protein; 423 AA.
XX
XX AC AAB30695;
XX
XX XX 02-APR-2001 (first entry)
XX DT
XX XX A fusion of anti-CD20 single chain antibody/streptavidin.
XX DE
XX XX Streptavidin; tumour cell; cancer; adenocarcinoma;
XX KW hematological malignancy; B9E9.
XX
XX OS Synthetic.
XX OS Streptomyces avidinii.
XX OS Homo sapiens.
XX
XX XX Key Location/Qualifiers
XX FH Protein 1..126
XX FT /note= "VH chain"
XX FT Peptide 127..150

```

```

FT Protein /note= "linker"
FT FT 151..259
FT FT /note= "VL chain"
FT FT 260..264
FT FT /note= "linker"
FT FT 265..423
FT FT /note= "streptavidin"
XX
XX PN WO200075333-A1.
XX
XX PD 14-DEC-2000.
XX
XX PF 05-JUN-2000; 2000WO-US15595.
XX PR 07-JUN-1999; 99US-0137900.
XX PR 03-DEC-1999; 99US-0168976.
XX PA (NEOR-) NEORX CORP.
XX PI Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;
XX PD WPI: 2001-091213/10.
XX DR N-PSDB; AAC86564.
XX
XX New vector constructs for expressing genomic streptavidin fusion
XX PT proteins which are useful for targeting tumour cells associated with
XX PT cancer, e.g. adenocarcinomas -
XX PA (NEOR-) NEORX CORP.
XX PI Example 2; Fig 11C; 100pp; English.
XX
XX The present sequence represents a fusion of an anti-CD20 single chain
XX antibody (B9E9) streptavidin. The fusion protein is expressed using
XX vectors of the invention. The specification describes vector constructs
XX for expressing streptavidin fusion proteins. The vector comprises a
XX nucleic acid encoding streptavidin or its functional variant operatively
XX linked to a promoter, and a cloning site for insertion of a second
XX nucleic acid sequence encoding a polypeptide to be fused with
XX streptavidin, interposed between the promoter and the first nucleic
XX acid sequence. Alternatively, the vector construct comprises a nucleic
XX acid, operatively linked to a promoter, encoding a polypeptide to be
XX fused with streptavidin, and a cloning site for insertion of a second
XX nucleic acid encoding at least 129 amino acids of streptavidin or its
XX functional variant. The fusion proteins are useful for targeting tumour
XX cells, particularly tumour cells associated with cancer,
XX e.g. adenocarcinomas or hematological malignancies. The vector construct
XX is useful for expressing of streptavidin fusion proteins. In particular,
XX these are useful as tools for medical diagnostics and therapeutic
XX purposes, e.g. for detecting the presence or absence of, or treating, a
XX target site within a mammalian host.
XX
XX Sequence 423 AA;
XX
XX Query Match 13.1%; Score 20; DB 22; Length 423;
XX Best Local Similarity 100.0%; Pred. No. 8.1e-10;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 49 TSYNHHWVKQTQGGLEWIG 68
XX 32 TSYNHHWVKQTQGGLEWIG 51
XX
XX Db
XX
XX RESULT 12
XX AAY84965
XX ID AAY84965 standard; Protein; 633 AA.
XX
XX AC AAY84965;
XX
XX XX 21-AUG-2000 (first entry)
XX DT
XX XX Amino acid sequence of a CD-20 specific chimeric receptor.
XX DE CD20-specific receptor; CD-20 specific redirected T cell; leukemia;
XX KW CD20+ malignancy; non-Hodgkin's lymphoma; myeloablative chemotherapy;
XX

```

KW stem cell rescue; autoimmune disease; lupus; rheumatoid arthritis.  
 XX Synthetic.  
 OS Mus sp.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 27..86 /note= "signal peptide from murine T86.66 antibody  
 FT kappa light chain"  
 FT Region 21..126  
 FT /note= "anti-CD20 variable regions"  
 FT Peptide 145..266 /note= "GS18 linker"  
 FT Region 283..392 /note= "hinge region"  
 FT Region 393..499 /note= "CH3 region"  
 FT Region 500..521  
 FT /note= "CD4 transmembrane region"  
 FT Region 522..633  
 FT /note= "zeta chain"  
 FN WO200023573-A2.  
 XX 27-APR-2000.  
 XX 20-OCT-1999; 99WO-US24484.  
 XX 20-OCT-1998; 98US-0105014.  
 XX (CITY ) CITY OF HOPE.  
 PA Raubitschek A, Jensen MC, Wu AM;  
 PI WPI; 2000-339676/29.  
 DR N-PSDB; AAL15019.  
 XX Genetically engineered CD20-specific redirected T cells useful for  
 FT treating a CD20+ malignancy, such as non-Hodgkin's lymphoma or CD20+  
 FT acute or chronic leukemia, and autoimmune disease -  
 XX Claim 10; Page 53-55; 58pp; English.  
 XX The present sequence represents a synthetic CD20-specific chimeric  
 CC receptor. The specification describes CD-20 specific redirected T cells  
 CC which express and bear on the cell surface membrane a CD20-chimeric  
 CC receptor comprising an intracellular signalling domain, a transmembrane  
 CC domain and an extracellular domain, the extracellular domain comprising  
 CC a CD20-specific receptor. The genetically engineered CD20-specific  
 CC redirected T cells are useful for treating a CD20+ malignancy, such  
 CC as non-Hodgkin's lymphoma or CD20+ acute or chronic leukemia, in a  
 CC human patient having previously undergone myeloablative chemotherapy and  
 CC stem cell rescue. The genetically engineered CD20-specific redirected  
 CC T cells are also useful for abrogating an untoward B cell function, such  
 CC as autoimmune disease (lupus or rheumatoid arthritis) in a patient.  
 XX  
 SQ Sequence 633 AA;  
 Query Match 13.1%; Score 20; DB 21; Length 633;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 49 TSYNMHWVKTPGGGLEWIG 68  
 Db 174 TSYNMHWVKTPGGGLEWIG 193  
 RESULT 13  
 AAR60565  
 ID AAR60565 standard; Protein: 124 AA.  
 XX  
 AC AAR60565;  
 XX

DT 25-MAR-2003 (updated)  
 DT 25-APR-1995 (first entry)  
 XX  
 DE Anti-carcinoembryonic antigen chimeric heavy chain Ab.  
 XX  
 KW Anti-carcinoembryonic antigen chimeric antibodies; CEAS;  
 KW chimeric human-murine; breast or colorectal carcinoma;  
 KW heavy chain.  
 XX  
 OS Chimeric Mus muscaris.  
 OS Chimeric Homo sapiens.  
 XX  
 PN WO9419466-A2.  
 XX  
 XX 01-SEP-1994.  
 XX 16-FEB-1994; 94WO-US01709.  
 XX 16-FEB-1993; 93US-0017570.  
 PR (DOWC ) DOW CHEM CO.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA Gourlie BB, Kaplan DA, Mezes PS, Rixon MW, Schlom J;  
 XX WPI; 1994-294331/36.  
 DR N-PSDB; AAO71395.  
 XX Anti-carcinoembryonic antigen chimeric antibodies - for diagnosis  
 PT and therapy of carcinoma, e.g. breast or colorectal carcinoma  
 XX  
 PS Claim 13; Page 51; 67pp; English.  
 XX  
 CC AAO71395 codes for AAR60565 the antibody heavy chain region of  
 CC murine-human anti-carcinoembryonic antigen (CEA) chimeric  
 CC antibody. Which can be used in in vitro immunoassays for  
 CC the detection of CEA, and monitoring of tumour-associated  
 CC antigen during therapy. It can also be used in vivo diagnostically,  
 CC or in therapy for the treatment of tumours associated with  
 CC colorectal and breast carcinomas, as well those of the  
 CC gastrointestinal tract, lung, ovary and pancreas.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 124 AA;  
 Query Match 12.4%; Score 19; DB 15; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 23 LQOSGAEIVRSGASVKMSC 41  
 Db 4 LQOSGAEIVRSGASVKMSC 22  
 RESULT 14  
 AAR60566  
 ID AAR60566 standard; Protein: 124 AA.  
 XX  
 XX AAR60566;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 25-APR-1995 (first entry)  
 XX  
 DE Anti-carcinoembryonic antigen chimeric heavy chain Ab.  
 XX  
 KW Anti-carcinoembryonic antigen chimeric antibodies; CEAS;  
 KW chimeric human-murine; breast or colorectal carcinoma;  
 KW heavy chain.  
 XX  
 OS Chimeric Mus muscaris.  
 OS Chimeric Homo sapiens.  
 XX  
 PN WO9419466-A2.

XX This sequence represents the murine COL1 VH chain (COL1muVH), and was  
 CC used to obtain the humanised antibody (Ab) of the invention. The  
 CC humanised Ab, or its fragment, specifically binds to carcinoembryonic  
 CC antigen (CEA). The Ab, optionally coupled to an effector or label, is  
 CC used to treat or prevent CEA-expressing cancers (e.g. of breast, ovary,  
 CC lung, stomach, or colon) and to detect CEA-expressing cells, either in  
 CC vitro (optionally on a solid support) or in vivo (particularly by tumour  
 CC imaging to identify tumours and metastases before surgery), for diagnosis  
 CC or prognosis. It is not significantly immunogenic, i.e. the Ab does not  
 CC induce a human anti-murine antibody or allergic response, or non-specific  
 CC cytotoxicity, so can be administered repeatedly. It retains specificity  
 CC for CEA, and has improved clearance (allowing efficient targeting) and  
 CC metabolic properties.  
 XX  
 SQ Sequence 124 AA;  
 Query Match 12.4%; Score 19; DB 20; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 23 LQSGAELVRSGASVRMSC 41  
 DB 4 LQSGAELVRSGASVRMSC 22  
 Search completed: August 30, 2003, 22:09:56  
 Job time : 28.1846 secs

XX 01-SEP-1994.  
 XX  
 XX 16-FEB-1994; 94WO-US01709.  
 XX  
 XX 16-FEB-1993; 93US-0017570.  
 XX  
 XX (DOWC ) DOW CHEM CO.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 XX Gourlie BB, Kaplan DA, Mezes PS, Rixon MW, Schlom J;  
 XX WPI; 1994-294331/36.  
 XX N-PSDB; AAQ71396.  
 XX  
 XX Anti-carcinoembryonic antigen chimeric antibodies - for diagnosis  
 XX and therapy of carcinoma, e.g. breast or colorectal carcinoma  
 XX  
 XX Claim 14; Page 53; 67pp; English.  
 XX  
 XX AAQ71396 codes for AAR60566 the antibody heavy chain region of  
 CC murine-human anti-carcinoembryonic antigen (CEA) chimeric  
 CC antibody. Which can be used in in vitro immunoassays for  
 CC the detection of CEA, and monitoring of tumour-associated  
 CC antigen during therapy. It can also be used in vivo diagnostically,  
 CC or in therapy for the treatment of tumours associated with  
 CC colorectal and breast carcinomas, as well those of the  
 CC gastrointestinal tract, lung, ovary and pancreas.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 124 AA;  
 Query Match 12.4%; Score 19; DB 15; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-09;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 23 LQSGAELVRSGASVRMSC 41  
 DB 4 LQSGAELVRSGASVRMSC 22  
 RESULT 15  
 AAY39528  
 ID AAY39528 standard; protein; 124 AA.  
 XX  
 XX AC AAY39528;  
 XX  
 XX DT 22-NOV-1999 (first entry)  
 XX  
 XX DE Murine COL1 VH chain (COL1muVH).  
 XX  
 XX KW Antibody; human; carcinoembryonic antigen; CEA; cancer; tumour imaging;  
 XX tumour identification; metastasis; diagnosis; mouse; VH chain.  
 XX  
 XX OS Mus sp.  
 XX  
 XX WO943817-A1.  
 XX  
 XX PN 02-SEP-1999.  
 XX  
 XX PF 25-FEB-1998; 98WO-US03680.  
 XX  
 XX PR 25-FEB-1998; 98WO-US03680.  
 XX  
 XX PA (DOWC ) DOW CHEM CO.  
 XX  
 XX Anderson WHK, Tempest PR, Carr FJ, Harris WJ, Armour K;  
 XX WPI; 1999-550870/46.  
 XX  
 XX PT New antibody specific for treatment and diagnosis of cancer -  
 XX Claim 2; Fig 1; 82pp; English.

**THIS PAGE BLANK (uspto)**



Qy 121 EGALDYGQGTSTVSSAKTTTPPVYPLVPGSL 153  
Db 121 EGALDYGQGTSTVSSAKTTTPPVYPLVPGSL 153

## RESULT 2

US-10-367-506-4  
; Sequence 4, Application US/10367506  
; Publication No. US20030152575A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; FILE OF INVENTION: TUMORS BEARING HMFG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/10/367,506  
; CURRENT FILING DATE: 2003-02-13  
; PRIOR APPLICATION NUMBER: US/09/861,294  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 153  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(19)  
US-10-367-506-4

Query Match 100.0%; Score 153; DB 12; Length 153;  
Best Local Similarity 100.0%; Pred. No. 5.3e-131;  
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECSWVFLFLISITGVHSGAYLQSGAELVRSGASVKMSCKASGYTLTSTNMHWKTP 60  
Db 1 MECSWVFLFLISITGVHSGAYLQSGAELVRSGASVKMSCKASGYTLTSTNMHWKTP 60  
Qy 61 GQGLEWIGNIPGNGDYYNKFQKASLTADTSSSTAYMQISLTSEDSAVYFCARGNW 120  
Db 61 GQGLEWIGNIPGNGDYYNKFQKASLTADTSSSTAYMQISLTSEDSAVYFCARGNW 120  
Qy 121 EGALDYGQGTSTVSSAKTTTPPVYPLVPGSL 153  
Db 121 EGALDYGQGTSTVSSAKTTTPPVYPLVPGSL 153

## RESULT 3

US-09-861-294-16  
; Sequence 16, Application US/09861294  
; Patent No. US20020098190A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; FILE OF INVENTION: TUMORS BEARING HMFG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/09/861,294  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 32  
; TYPE: PRT

; ORGANISM: Mus musculus  
US-09-861-294-16

Query Match 20.9%; Score 32; DB 9; Length 32;  
Best Local Similarity 100.0%; Pred. No. 6e-22;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 KASLTADTSSSTAYMQISLTSEDSAVYFCAR 117  
Db 1 KASLTADTSSSTAYMQISLTSEDSAVYFCAR 32

## RESULT 4

US-10-367-506-16  
; Sequence 16, Application US/10367506  
; Publication No. US20030152575A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; FILE OF INVENTION: TUMORS BEARING HMFG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/10/367,506  
; CURRENT FILING DATE: 2003-02-13  
; PRIOR APPLICATION NUMBER: US/09/861,294  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 32  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-367-506-16

Query Match 20.9%; Score 32; DB 12; Length 32;  
Best Local Similarity 100.0%; Pred. No. 6e-22;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 KASLTADTSSSTAYMQISLTSEDSAVYFCAR 117  
Db 1 KASLTADTSSSTAYMQISLTSEDSAVYFCAR 32

## RESULT 5

US-09-990-205-4  
; Sequence 4, Application US/09990205  
; Patent No. US20020150572A1  
; GENERAL INFORMATION:  
; APPLICANT: FOON, Kenneth A.  
; APPLICANT: CHATTERJEE, Malaya  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF PSORIASIS  
; FILE REFERENCE: 304142000501  
; CURRENT APPLICATION NUMBER: US/09/990,205  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: U.S. 09/192,838  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: U.S. 60/065,774  
; PRIOR FILING DATE: 1997-11-17  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: Mus Musculus  
US-09-990-205-4

Query Match 20.3%; Score 31; DB 10; Length 152;  
Best Local Similarity 100.0%; Pred. No. 1.8e-20;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 123 ALDWGQGTSTVSSAKTTPPPYPLVPGSL 153  
|||||  
Db 122 ALDWGQGTSTVSSAKTTPPPYPLVPGSL 152

## RESULT 6

US-10-153-401-4  
; Sequence 4, Application US/10153401  
; Publication No. US20030114398A1  
; GENERAL INFORMATION:  
; APPLICANT: Chatterjee, Malaya  
; Foon, Kenneth A.  
; Chatterjee, Sunil K.

; TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE  
; TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA

; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: PALO ALTO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/153,401  
; FILING DATE: 27-Aug-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/293,533  
; FILING DATE: 1999-04-15  
; APPLICATION NUMBER: US 08/372,676  
; FILING DATE: 1995-01-17  
; APPLICATION NUMBER: US 08/591,196  
; FILING DATE: 1996-01-16  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Catherine M. Polizzi  
; REGISTRATION NUMBER: 40,130  
; REFERENCE/DOCKET NUMBER: 304142000202  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 813-5600  
; TELEFAX: (415) 494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 152 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-153-401-4

Query Match 20.3%; Score 31; DB 15; Length 152;  
Best Local Similarity 100.0%; Pred. No. 1.8e-20;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 ALDWGQGTSTVSSAKTTPPPYPLVPGSL 153  
|||||  
Db 122 ALDWGQGTSTVSSAKTTPPPYPLVPGSL 152

## RESULT 7

US-09-861-294-12  
; Sequence 12, Application US/09861294  
; Patent No. US20020098190A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TUMORS BEARING HMFG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/09/861,294  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-861-294-12

Query Match 19.6%; Score 30; DB 9; Length 30;  
Best Local Similarity 100.0%; Pred. No. 3.7e-20;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QAYLQSGAELVRSGASVKMSCKASGYTLT 49  
|||||  
Db 1 QAYLQSGAELVRSGASVKMSCKASGYTLT 30

## RESULT 8

US-10-367-506-12  
; Sequence 12, Application US/10367506  
; Publication No. US20030152575A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TUMORS BEARING HMFG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/10/367,506  
; CURRENT FILING DATE: 2003-02-13  
; PRIOR APPLICATION NUMBER: US/09/861,294  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-367-506-12

Query Match 19.6%; Score 30; DB 12; Length 30;  
Best Local Similarity 100.0%; Pred. No. 3.7e-20;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QAYLQSGAELVRSGASVKMSCKASGYTLT 49  
|||||  
Db 1 QAYLQSGAELVRSGASVKMSCKASGYTLT 30

## RESULT 9

US-10-244-821-6  
; Sequence 6, Application US/10244821  
; Publication No. US20030143233A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen Charles  
; APPLICANT: Graves, Scott Stoll  
; APPLICANT: Schultz, Joanne Elaine  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James Allen  
; APPLICANT: Reno, John M.  
; APPLICANT: Dearstyne, Erica A.

```

; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted amino acid sequence of B9E9 s
; OTHER INFORMATION: antibody-genomic streptavidin fusion
US-10-244-821-6

```

Query Match	13.1%	Score 20;	DB 12;	Length 412;
Best Local Similarity	100.0%;	Pred. No. 3.9e-10;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

RESULT 10
US-10-013-173-6
; Sequence 6, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, Jonh M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted amino acid sequence of B9E9 single chain
; US-10-013-173-6
; OTHER INFORMATION: antibody-genomic streptavidin fusion

```

Query Match	13.1%	Score 20;	DB 15;	Length 412;
Best Local Similarity	100.0%;	Pred. No. 3.9e-10;		
Matches 20;	Conservative	0;	Mismatches 0;	Indels 0;
				Gaps 0;

RESULT 11  
US-10-150-762-6  
; Sequence 6, Application US/10150762  
; Publication No. US20030103948A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen C.  
; APPLICANT: Graves, Scott S.  
; APPLICANT: Schwartz, Joanne B.  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James A.  
; APPLICANT: Reno, Jonh M.  
; APPLICANT: Dearnstyne, Erica A.

```

; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
;
; TITLE OF INVENTION: METHODS OF USE THEREOF
;
; FILE REFERENCE: 690022.547C2
;
; CURRENT APPLICATION NUMBER: US/10/150,762
;
; CURRENT FILING DATE: 2002-05-17
;
; NUMBER OF SEQ ID NOS: 90
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 6
;
; LENGTH: 412
;
; TYPE: PRT
;
; ORGANISM: Artificial Sequence
;
; FEATURE:
;
; OTHER INFORMATION: Predicted amino acid sequence of B9E9 s
;
; OTHER INFORMATION: antibody-genomic streptavidin fusion
;
; US-10-150-762-6

```

Query Match	13.1%	Score 20;	DB 15;	Length 412;
Best Local Similarity	100.0%;	Pred. No. 3.9e-10;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

RESULT 12
US/10/244
; Sequence 8, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyn, Erica A.
; TITLE OF INVENTION: STRETAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244, 821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted amino acid sequence of B9E9 s
; US/10/244, 821-8

```

Query Match	13.1%	Score 20;	DB 12;	Length 423;
Best Local Similarity	100.0%;	Pred. No. 3.9e-10;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

RESULT 13
US/10/013
; Sequence 8, Application US/10013173
; Publication NO. US2003009597A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, Jonh M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND

```

; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 690022.547C1  
; CURRENT APPLICATION NUMBER: US/10/013,173  
; CURRENT FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 423  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Predicted amino acid sequence of B9E9 single chain antibody-  
US/10/013,173-8

Query Match 13.1%; Score 20; DB 15; Length 423;  
Best Local Similarity 100.0%; Pred. No. 3.9e-10; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0;

QY 49 TSYNMHWVKQTGGGLEWIG 68  
Db 32 TSYNMHWVKQTGGGLEWIG 51  
|||||

## RESULT 14

US/10/150  
; Sequence 8, Application US/10150762  
; Publication No. US20030103948A1

; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen C.  
; APPLICANT: Graves, Scott S.  
; APPLICANT: Schultz, Joanne E.  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James A.  
; APPLICANT: Reno, John M.  
; APPLICANT: Dearstyne, Erica A.

; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
; FILE REFERENCE: 690022.547C2

; CURRENT APPLICATION NUMBER: US/10/150,762  
; CURRENT FILING DATE: 2002-05-17  
; NUMBER OF SEQ ID NOS: 90

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8

; LENGTH: 423  
; TYPE: PRT

; ORGANISM: Artificial Sequence  
; FEATURE:

; OTHER INFORMATION: Predicted amino acid sequence of B9E9 single chain antibody-  
US/10/150,762-8

Query Match 13.1%; Score 20; DB 15; Length 423;  
Best Local Similarity 100.0%; Pred. No. 3.9e-10; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0;

QY 49 TSYNMHWVKQTGGGLEWIG 68  
Db 32 TSYNMHWVKQTGGGLEWIG 51  
|||||

## RESULT 15

US-10-194-975-121

; Sequence 121, Application US/10194975

; Publication No. US20030039649A1

; GENERAL INFORMATION:

; APPLICANT: Foote, Jefferson  
; TITLE OF INVENTION: Super Humanized Antibodies

; FILE REFERENCE: 501231.01

; CURRENT APPLICATION NUMBER: US/10/194,975

; CURRENT FILING DATE: 2002-10-10

; PRIOR APPLICATION NUMBER: US 60/305,111

; PRIOR FILING DATE: 2001-07-12

; NUMBER OF SEQ ID NOS: 122

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 121  
; LENGTH: 119  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-10-194-975-121

Query Match 12.4%; Score 19; DB 15; Length 119;  
Best Local Similarity 100.0%; Pred. No. 1.1e-09; Mismatches 0; Indels 0; Gaps 0;  
Matches 19; Conservative 0;

QY 50 SYNMHWVKQTGGGLEWIG 68  
Db 31 SYNMHWVKQTGGGLEWIG 49  
|||||

Search completed: August 30, 2003, 22:12:33  
Job time : 17.4295 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 22:07:01 ; Search time 10.7819 seconds  
(without alignments)  
1364.679 Million cell updates/sec

Title: US-08-836-455-4

Perfect score: 153

Sequence: 1 MGSWVFLFLSITTCVHSQ.....TVSSAKTTPPVPLVPGSL 153

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_76:\*\*

1: pir1:\*\*

2: pir2:\*\*

3: pir3:\*\*

4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	14.4	119	2	PL0089
2	22	14.4	151	2	PL0011
3	20	13.1	93	2	S42182
4	20	13.1	94	2	S42177
5	20	13.1	95	2	S42178
6	20	13.1	97	2	S42181
7	20	13.1	101	2	S42179
8	20	13.1	101	2	S42184
9	20	13.1	102	2	S42180
10	19	12.4	140	2	S14238
11	19	12.4	149	2	S30752
12	19	12.4	150	2	PN0444
13	19	12.4	152	2	B26471
14	19	12.4	178	2	S29594
15	18	11.8	124	2	S06824
16	17	11.1	112	2	S26473
17	17	11.1	246	2	S38950
18	17	11.1	446	2	S40295
19	16	10.5	109	2	PH1001
20	16	10.5	111	2	PH0998
21	16	10.5	116	2	S09962
22	16	10.5	119	2	PH1512
23	16	10.5	140	2	S09216
24	15	9.8	43	2	S11111
25	15	9.8	67	2	H28833
26	15	9.8	91	2	PL0242
27	15	9.8	94	2	G32513
28	15	9.8	96	2	H28195
29	15	9.8	106	2	PH1002

30 15 9.8 107 2 PL0240 Ig heavy chain V r  
31 15 9.8 107 2 PL0241 Ig heavy chain V r  
32 15 9.8 108 2 S26316 Ig heavy chain V r  
33 15 9.8 109 2 PH0997 Ig heavy chain V r  
34 15 9.8 110 2 S26317 Ig heavy chain V r  
35 15 9.8 110 2 PH1000 Ig heavy chain V r  
36 15 9.8 111 2 S25033 Ig heavy chain V r  
37 15 9.8 111 2 S25034 Ig heavy chain V r  
38 15 9.8 112 2 S09957 Ig heavy chain V-D  
39 15 9.8 113 2 S55535 Ig heavy chain V r  
40 15 9.8 115 2 PL0238 Ig heavy chain V r  
41 15 9.8 115 2 A54378 Ig heavy chain V r  
42 15 9.8 117 2 PL0237 Ig heavy chain V r  
43 15 9.8 117 2 PL0235 Ig heavy chain V r  
44 15 9.8 117 2 PL0234 Ig heavy chain V r  
45 15 9.8 117 2 S19966 Ig heavy chain V r

#### ALIGNMENTS

##### RESULT 1

PL0089  
Ig heavy chain V region (12sl8-1) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 04-Mar-2000  
C:Accession: PL0089  
R:Meek, K.; Hasemann, C.; Pollok, B.; Alkan, S.S.; Brait, M.; Siaoui, M.; Urbain, J.;  
J. Exp. Med. 169, 519-533, 1989  
A:Title: Structural characterization of antidiotypic antibodies; evidence that Ab2s  
A:Reference number: PL0080; MUID:89094248; PMID:2492056  
A:Accession: PL0089  
A:Molecule type: mRNA  
A:Residues: 1-119 <MEE>  
A:Cross-references: GB:X58580; GB:Y00794; NID:951591; PIDN:CAA1456.1; PID:9930150  
A:Note: the sequence shown here is from the VH region of an antidiotypic monoclonal  
A:Note: sequences from two other clones (18S28-16 and 12S84-3) were almost identical  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 14.4%; Score 22; DB 2; Length 119;

Best Local Similarity 100.0%; Pred. No. 1.4e-14;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 DTSSSTAYMOISLTSEDSAVY 113

Db 73 DTSSSTAYMOISLTSEDSAVY 94

##### RESULT 2

PL0011

Ig heavy chain precursor V region (4C11) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 21-Jan-2000

C:Accession: PL0011

R:Cheng, H.L.; Sood, A.K.; Ward, R.E.; Kleber-Emmons, T.; Kohler, H.

Mol. Immunol. 25, 33-40, 1988

A:Title: Structural basis of stimulatory anti-idiotypic antibodies.

A:Reference number: PL0011; MUID:88142863; PMID:3125424

A:Accession: PL0011

A:Molecule type: mRNA

A:Residues: 1-151 <CHE>

A:Experimental source: cell line 4C11

C:Comment: This protein is an anti-idiotypic antibody that induces an anti-phosphoryl

C:Superfamily: immunoglobulin V region; immunoglobulin

C:Keywords: heterotrimer; immunoglobulin

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-136/Product: Ig heavy chain V region 4C11 #status predicted <NAT>

F:34-117/Domain: immunoglobulin homology <IMM>

F:50-54/Region: complementarity-determining 1

F:69-85/Region: complementarity-determining 2

F:118-125/Region: complementarity-determining 3

F:137-151/Domain: C region (fragment) #status predicted <COR>

Query Match 14.4%; Score 22; DB 2; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1.7e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 GADYWGQGTSTVTSAAKTPP 143  
|||||  
Db 121 GADYWGQGTSTVTSAAKTPP 142  
|||||

## RESULT 3

S42182

Ig gamma chain V region - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Sep-1994 #sequence\_revision 21-Jul-1995 #text\_change 23-Jul-1999  
C:Accession: S42182  
R:Mo, J.A.; Bona, C.A.; Holmdahl, R.  
Eur. J. Immunol. 23, 2503-2510, 1993  
A:Title: Variable region gene selection of immunoglobulin G-expressing B cells with spec  
A:Reference number: S42176; MUID:94009207; PMID:7691608  
A:Accession: S42182  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-93 <MOJ>  
A:Cross-references: EMBL:Z25453; NID:g407822; PIDN:CAA80940.1; PID:g407823  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:8-91/Domain: immunoglobulin homology <IMM>

Query Match 13.1%; Score 20; DB 2; Length 93;  
Best Local Similarity 100.0%; Pred. No. 1.1e-12;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 TSYNMHWVKTPGGGLEWIG 68  
|||||  
Db 23 TSYNMHWVKTPGGGLEWIG 42  
|||||

## RESULT 4

S42177

Ig gamma chain V region - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Sep-1994 #sequence\_revision 21-Jul-1995 #text\_change 23-Jul-1999  
C:Accession: S42177  
R:Mo, J.A.; Bona, C.A.; Holmdahl, R.  
Eur. J. Immunol. 23, 2503-2510, 1993  
A:Title: Variable region gene selection of immunoglobulin G-expressing B cells with spec  
A:Reference number: S42176; MUID:94009207; PMID:7691608  
A:Accession: S42177  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-94 <MOJ>  
A:Cross-references: EMBL:Z25443; NID:g407812; PIDN:CAA80930.1; PID:g407813  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:8-91/Domain: immunoglobulin homology <IMM>

Query Match 13.1%; Score 20; DB 2; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 TSYNMHWVKTPGGGLEWIG 68  
|||||  
Db 23 TSYNMHWVKTPGGGLEWIG 42  
|||||

## RESULT 5

S42178

Ig gamma chain V region - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Sep-1994 #sequence\_revision 21-Jul-1995 #text\_change 23-Jul-1999  
C:Accession: S42178

R:Mo, J.A.; Bona, C.A.; Holmdahl, R.  
Eur. J. Immunol. 23, 2503-2510, 1993  
A:Title: Variable region gene selection of immunoglobulin G-expressing B cells with s  
A:Reference number: S42176; MUID:94009207; PMID:7691608  
A:Accession: S42178  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-95 <MOJ>  
A:Cross-references: EMBL:Z25445; NID:g407814; PIDN:CAA80932.1; PID:g407815  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:8-91/Domain: immunoglobulin homology <IMM>

Query Match 13.1%; Score 20; DB 2; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 TSYNMHWVKTPGGGLEWIG 68  
|||||  
Db 23 TSYNMHWVKTPGGGLEWIG 42  
|||||

## RESULT 6

S42181

Ig gamma chain V region - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Sep-1994 #sequence\_revision 21-Jul-1995 #text\_change 23-Jul-1999  
C:Accession: S42181  
R:Mo, J.A.; Bona, C.A.; Holmdahl, R.  
Eur. J. Immunol. 23, 2503-2510, 1993  
A:Title: Variable region gene selection of immunoglobulin G-expressing B cells with s  
A:Reference number: S42176; MUID:94009207; PMID:7691608  
A:Accession: S42181  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-97 <MOJ>  
A:Cross-references: EMBL:Z25451; NID:g407820; PIDN:CAA80938.1; PID:g407821  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:8-91/Domain: immunoglobulin homology <IMM>

Query Match 13.1%; Score 20; DB 2; Length 97;  
Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 TSYNMHWVKTPGGGLEWIG 68  
|||||  
Db 23 TSYNMHWVKTPGGGLEWIG 42  
|||||

## RESULT 7

S42179

Ig gamma chain V region - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Sep-1994 #sequence\_revision 21-Jul-1995 #text\_change 23-Jul-1999  
C:Accession: S42179  
R:Mo, J.A.; Bona, C.A.; Holmdahl, R.  
Eur. J. Immunol. 23, 2503-2510, 1993  
A:Title: Variable region gene selection of immunoglobulin G-expressing B cells with s  
A:Reference number: S42176; MUID:94009207; PMID:7691608  
A:Accession: S42179  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-101 <MOJ>  
A:Cross-references: EMBL:Z25447; NID:g407816; PIDN:CAA80934.1; PID:g407817  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 13.1%; Score 20; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 TSYNNHWWKQPPGGGLEWIG 68  
 |||||  
 Db 30 TSYNNHWWKQPPGGGLEWIG 49

## RESULT 8

S42184  
 Ig gamma chain V region - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 28-Sep-1994 #sequence\_revision 21-Jul-1995 #text\_change 06-Jun-1997  
 C:Accession: S42184  
 R:Mo, J.A.; Bona, C.A.; Holmdahl, R.  
 Eur. J. Immunol. 23, 2503-2510, 1993  
 A:Title: Variable region gene selection of immunoglobulin G-expressing B cells with spec  
 A:Reference number: S42176; MUID:94009207; PMID:7691608  
 A:Accession: S42184  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-101 <MO>  
 A:Cross-references: EMBL:Z25457  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 13.1%; Score 20; DB 2; Length 101;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 TSYNNHWWKQPPGGGLEWIG 68  
 |||||  
 Db 30 TSYNNHWWKQPPGGGLEWIG 49

## RESULT 9

S42180  
 Ig gamma chain V region - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 28-Sep-1994 #sequence\_revision 21-Jul-1995 #text\_change 23-Jul-1999  
 C:Accession: S42180  
 R:Mo, J.A.; Bona, C.A.; Holmdahl, R.  
 Eur. J. Immunol. 23, 2503-2510, 1993  
 A:Title: Variable region gene selection of immunoglobulin G-expressing B cells with spec  
 A:Reference number: S42176; MUID:94009207; PMID:7691608  
 A:Accession: S42180  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-102 <MO>  
 A:Cross-references: EMBL:Z25449; NID:g407818; PIDN:CAA80936.1; PID:g407819  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 13.1%; Score 20; DB 2; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 TSYNNHWWKQPPGGGLEWIG 68  
 |||||  
 Db 30 TSYNNHWWKQPPGGGLEWIG 49

## RESULT 10

S14238  
 Ig gamma-1 chain precursor (15C5) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 25-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999  
 C:Accession: S14238  
 R:Vandamme, A.M.; Bulens, F.; Bernar, H.; Nelles, L.; Lijnen, R.H.; Collen, D.  
 Eur. J. Biochem. 192, 767-775, 1990  
 A:Title: Construction and characterization of a recombinant murine monoclonal antibody c  
 A:Reference number: S14236; MUID:91006173; PMID:2209622  
 A:Accession: S14238  
 A:Molecule type: mRNA

A:Residues: 1-140 <VAN>  
 A:Cross-references: EMBL:X56392; NID:g51619; PIDN:CAA39803.1; PID:g747853  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:26-108/Domain: immunoglobulin homology <IMM>

Query Match 12.4%; Score 19; DB 2; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 DYWGQGTSTVTSSAKTTTP 143  
 |||||  
 Db 116 DYWGQGTSTVTSSAKTTTP 134

## RESULT 11

S30752  
 Ig heavy chain precursor V region - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 23-Jul-1999  
 C:Accession: S30752  
 R:Grant, F.J.; Levin, S.D.; Gilbert, T.; Kindsvogel, W.  
 Nucleic Acids Res. 15, 5496, 1987  
 A:Title: Improved RNA sequencing method to determine immunoglobulin mRNA sequence.  
 A:Reference number: S30751; MUID:87260030; PMID:3601683  
 A:Accession: S30752  
 A:Molecule type: mRNA  
 A:Residues: 1-149 <GRA>  
 A:Cross-references: EMBL:X05878; NID:g52526; PIDN:CAA29302.1; PID:g52527  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:33-116/Domain: immunoglobulin homology <IMM>  
 F:138-149/Domain: C region (C-gamma 2b) (fragment) #status predicted <CRE>

Query Match 12.4%; Score 19; DB 2; Length 149;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-11;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 DYWGQGTSTVTSSAKTTTP 143  
 |||||  
 Db 125 DYWGQGTSTVTSSAKTTTP 143

## RESULT 12

PN0444  
 Ig heavy chain V region precursor - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000  
 C:Accession: PN0444  
 R:Kaluza, B.; Betzel, G.; Shao, H.; Diamantstein, T.; Weidle, U.H.  
 Gene 122, 321-328, 1992  
 A:Title: A general method for chimerization of monoclonal antibodies by inverse polym  
 A:Reference number: PN0444; MUID:93138402; PMID:1339379  
 A:Accession: PN0444  
 A:Molecule type: mRNA  
 A:Residues: 1-150 <KAL>

A:Cross-references: GB:L02346  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-150/Product: Ig heavy chain V region #status predicted <MAT>  
 F:20-117/Domain: variable region <VRG>  
 F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 12.4%; Score 19; DB 2; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-11;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 DYWGQGTSTVTSSAKTTTP 143  
 |||||  
 Db 125 DYWGQGTSTVTSSAKTTTP 143

## RESULT 13

B26471  
Ig heavy chain precursor V region (MAK33) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 05-Jun-1988 #sequence\_revision 05-Jun-1988 #text\_change 23-Jul-1999  
R:Accession: B26471; S70410  
C:Buckel, P.; Hubner-Parajsz, C.; Mattes, R.; Lenz, H.; Haug, H.; Beaucamp, K.  
Gene 51, 13-19, 1987  
A:Title: Cloning and nucleotide sequence of heavy- and light-chain cDNAs from a creatine  
A:Reference number: A91572; MUID:87248058; PMID:3110009  
A:Accession: B26471  
A:Molecule type: mRNA  
A:Residues: 1-152 <BUC>  
A:Cross-references: GB:M1613; NID:g195405; PIDN:AAA38292.1; PID:g195406  
R:Lebecque, S.G.; Gearhart, P.J.  
J. Exp. Med. 172, 1717-1727, 1990  
A:Title: Boundaries of somatic mutation in rearranged immunoglobulin genes: 5' boundary  
A:Reference number: S70410; MUID:91079775; PMID:2258702  
A:Accession: S70410  
A>Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-19 <LEB>  
A:Cross-references: EMBL:X53776; NID:g52475; PIDN:CAA37792.1; PID:g52476  
C:Genetics:  
A:Introns: 16/1  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-152/Product: Ig heavy chain V region MAK33 #status predicted <MAT>  
F;34-117/Domain: immunoglobulin homology <IMM>

Query Match 12.4%; Score 19; DB 2; Length 152;

Best Local Similarity 100.0%; Pred. No. 1.8e-11;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 DYWGQGTSTVTSSAKTTP 143

|||||

DB 130 DYWGQGTSTVTSSAKTTP 148

## RESULT 14

S29594  
Ig gamma chain (WM65) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999  
C:Accession: S29594  
R:Seymour, R.  
submitted to the EMBL Data Library, February 1991  
A:Reference number: S29593  
A:Accession: S29594  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-178 <SEY>  
A:Cross-references: EMBL:X57857; NID:g52590; PIDN:CAA40992.1; PID:g52591  
C:Keywords: immunoglobulin

Query Match

Best Local Similarity 100.0%; Pred. No. 2.1e-11;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 DYWGQGTSTVTSSAKTTP 143

|||||

DB 119 DYWGQGTSTVTSSAKTTP 137

## RESULT 15

S06824  
Ig heavy chain V region (clone 12D4) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1991 #sequence\_revision 11-Nov-1994 #text\_change 23-Jul-1999  
C:Accession: S06824  
R:Miller III, A.; Glasel, J.A.  
J. Mol. Biol. 209, 763-778, 1989

A:Title: Comparative sequence and immunochemical analyses of murine monoclonal anti-m  
A:Reference number: S06815; MUID:90064531; PMID:2555519  
A:Accession: S06824  
A:Molecule type: mRNA  
A:Residues: 1-124 <MIL>  
A:Cross-references: EMBL:X17166; NID:g51918; PIDN:CAA35044.1; PID:g930156  
A:Note: the authors did not translate the codon TGG for residue 36  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F;15-98/Domain: immunoglobulin homology <IMM>  
F;22-96/Disulfide bonds: #status predicted

Query Match 11.8%; Score 18; DB 2; Length 124;

Best Local Similarity 100.0%; Pred. No. 1.5e-10;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 DYWGQGTSTVTSSAKTTP 142

|||||

DB 107 DYWGQGTSTVTSSAKTTP 124

Search completed: August 30, 2003, 22:11:49

Job time : 10.7819 secs





OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=80078170; PubMed=6765983;  
 RA Schilling J., Clevinger B., Davie J.M., Hood L.;  
 RT "Amino acid sequence of homogeneous antibodies to dextran and DNA  
 rearrangements in heavy chain V-region gene segments.";  
 RL Nature 283:35-40(1980).  
 CC -1- MISCELLANEOUS: THE SEQUENCES OF 10 HYBRIDOMA PROTEINS THAT ALSO  
 BIND DEXTRAN DIFFER FROM THAT SHOWN AT 1-7 POSITIONS, MANY OF  
 WHICH OCCUR IN THE D AND J SEGMENTS.  
 CC -1- MISCELLANEOUS: THIS PROTEIN BINDS DEXTRAN.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 DR PIR: A26242; MHMSJ5.  
 DR HSSP; P01789; IMCP.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG-LIKE; 1.  
 DR Immunoglobulin v region.  
 KW DOMAIN 1 116  
 FT DISULFID 22 96  
 FT NON\_TER 117 117  
 SQ SEQUENCE 117 AA; 13024 MW; 292E2AF4BE447E41 CRC64;  
 Query Match 9.2%; Score 14; DB 1; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 5e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 34 GASVKMSCKASGYT 47  
 DB 15 GASVKMSCKASGYT 28  
 RESULT 3  
 HV03\_MOUSE  
 ID HV03\_MOUSE STANDARD; PRT; 120 AA.  
 AC P01747;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ig heavy chain V region 36-65.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83131846; PubMed=6186498;  
 RA Siekevitz M., Geffer M.L., Brodeur P., Riblet R.,  
 RA Marshak-Rothstein A.;  
 RT "The genetic basis of antibody production: the dominant anti-arsenate  
 idiotypic response of the strain A mouse.";  
 RL Eur. J. Immunol. 12:1023-1032(1982).  
 CC -1- MISCELLANEOUS: FROM ANALYSIS OF THE SIZES OF SEVERAL OTHER  
 DIFFERENTIATED GENES THAT HYBRIDIZE TO THIS ONE, THE AUTHORS  
 CONCLUDE THAT ALL OF THESE V REGIONS HAVE REARRANGED TO THE SAME J  
 SEGMENT, JH2.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 DR HSSP; P01789; IMCP.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG-LIKE; 1.  
 KW Immunoglobulin v region; Hybridoma.  
 FT DOMAIN 1 111  
 FT IG-LIKE.

FT NON\_TER 120  
 SQ SEQUENCE 120 AA; 13307 MW; FF04E4A167B654AF CRC64;  
 Query Match 9.2%; Score 14; DB 1; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 104 SLTSEDSAVYFCAR 117  
 DB 84 SLTSEDSAVYFCAR 97  
 RESULT 4  
 HV02\_MOUSE  
 ID HV02\_MOUSE STANDARD; PRT; 140 AA.  
 AC P01746;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ig heavy chain V region 93G7 precursor.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-A/J;  
 RX MEDLINE=82152818; PubMed=6801765;  
 RA Sims J., Rabbitts T.H., Estess P., Slaughter C., Tucker P.W.,  
 RA Capra J.D.;  
 RT "Somatic mutation in genes for the variable portion of the  
 immunoglobulin heavy chain.";  
 RL Science 216:309-311(1982).  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; J00493; AAA38128.1;  
 CC PIR; A94264; HVNSG7.  
 DR HSSP; P01810; 2FBJ.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG-LIKE; 1.  
 KW Immunoglobulin v region; Hybridoma; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 140 IG HEAVY CHAIN V REGION 93G7.  
 FT DOMAIN 20 139 IG-LIKE.  
 FT NON\_TER 140 140  
 SQ SEQUENCE 140 AA; 15514 MW; 25A4CBBE31DA5CE8 CRC64;  
 Query Match 9.2%; Score 14; DB 1; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 104 SLTSEDSAVYFCAR 117  
 DB 104 SLTSEDSAVYFCAR 117  
 RESULT 5  
 HV43\_MOUSE  
 ID HV43\_MOUSE STANDARD; PRT; 144 AA.  
 AC P01819;  
 DT 21-JUL-1986 (Rel. 01, Created)

```
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DI Ig heavy chain V region MOPC 141 precursor.
OS Mus musculus (Mouse).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81012133; PubMed=6774258;
RA Sakano H., Maki R., Kurosawa Y., Roeder W., Tonegawa S.;
RT "Two types of somatic recombination are necessary for the generation
RT of complete immunoglobulin heavy-chain genes.";
RL Nature 286:676-683(1990).
CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS TRANSLATED FROM A
CC DIFFERENTIATED GENE ISOLATED FROM A MYELOMA THAT SECRETES IGG2B.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00768; CAA24149.1; -.
DR PIR; A02094; G2MS14.
DR HSP; P01825; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 144 IG HEAVY CHAIN V REGION MOPC 141.
FT DOMAIN 20 130 IG-LIKE.
FT NON_TER 144 144
FT SEQUENCE 144 AA; 15759 MW; 8E47A7CB3706D30A CRC64;

Query Match 9.2%; Score 14; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 6.1e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LDYWGQGTSTVSS 137
DB 131 LDYWGQGTSTVSS 144
|||||
RESULT 6
HV16_MOUSE STANDARD; PRT; 136 AA.
AC P01783;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Ig heavy chain V region MOPC 21 precursor (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Rech M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPb family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
RN [2]
RP SEQUENCE OF 17-136.
```

```
RX MEDLINE=77100368; PubMed=401950;
RA Adetugbo K., Milstein C., Secher D.S.;
RT "Molecular analysis of spontaneous somatic mutants.";
RL Nature 265:299-304(1977).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00522; AAD15290.1; -.
DR PIR; E90809; GIMS21.
DR PDB; 1IGC; O3-JUN-95.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal; 3D-structure.
FT NON_TER 1 16
FT SIGNAL <1 16
FT CHAIN 17 136 IG HEAVY CHAIN V REGION MOPC 21.
FT DOMAIN 115 119 D SEGMENT.
FT DOMAIN 120 136 JH4 SEGMENT.
FT DISULFID 38 112
FT CONFLICT 75 78 HYAD -> DYAH (IN REF. 2).
FT CONFLICT 89 90 DN -> ND (IN REF. 2).
FT CONFLICT 115 115 W -> H (IN REF. 2).
FT CONFLICT 120 120 Y -> W (IN REF. 2).
FT NON_TER 136 136
FT SEQUENCE 136 AA; 15071 MW; 2276A98DBDF7016 CRC64;

Query Match 8.5%; Score 13; DB 1; Length 136;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 DYWGQGTSTVSS 137
DB 124 DYWGQGTSTVSS 136
|||||
RESULT 7
HV42_MOUSE STANDARD; PRT; 117 AA.
AC P01812;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region MOPC 173.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-104.
RX MEDLINE=72105531; PubMed=5062012;
RA Bourgois A., Fougereau M., de Preval C.;
RT "Sequence of amino acids of the NH 2 -terminal region of a
RT mouse-clonal immunoglobulin heavy chain.";
RL Eur. J. Biochem. 24:446-455(1972).
RN [2]
RP SEQUENCE OF 105-117.
RX MEDLINE=76091933; PubMed=812695;
RA Rocca-Serra J., Milili M., Fougereau M.;
RT "Determination of the primary structure of a mouse IgG2a
RT immunoglobulin. Amino-acid sequence of the H4 cyanogen-bromide
RT fragment.";
RL Eur. J. Biochem. 59:511-523(1975).
RN [3]
```

RP SEQUENCE OF 96-117 FROM N.A.  
RX MEDLINE-81223769; PubMed-6787590;  
RA Gough N.M., Bernard O.;  
RT "Sequences of the joining region genes for immunoglobulin heavy  
chains and their role in generation of antibody diversity.";  
RL Proc. Natl. Acad. Sci. U.S.A. 78:509-513(1981).  
RN [4]  
RP DISULFIDE BOND.  
RA Bourgois A., Fougereau M.;  
RT "Partial amino acid sequence of the variable region of a mouse  
gamma2a immunoglobulin heavy chain. Evidence for the existence of a  
third sub-group of variability for the heavy chain pool.";  
RL FEBS Lett. 8:263-268(1970).  
CC -1- MISCELLANEOUS: THIS GAMMA-2A CHAIN WAS ISOLATED FROM A MYELOMA  
PROTEIN.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.  
PIR: A91190; G2MS73.  
DR HSSP; P01810; 2FBJ.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS50835; IG-LIKE; 1.  
KW Immunoglobulin V region.  
FT DOMAIN 1 116 IG-LIKE.  
FT DISULFID 22 96  
FT CONFLICT 105 105 N -> D (IN REF. 2).  
FT NON\_TER 117 117  
SQ SEQUENCE 117 AA; 13051 MW; 156DCCC259380F19 CRC64;  
Query Match 7.8%; Score 12; DB 1; Length 117;  
Best Local Similarity 100.0%; Pred. No. 5.6e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 126 YWGQSTVTSS 137  
DB 106 YWGQSTVTSS 117  
RESULT 8  
ID HV04\_MOUSE STANDARD; PRT; 117 AA.  
AC P01748;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V region 23 precursor.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6;  
RX MEDLINE-81234548; PubMed-6788376;  
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,  
RA Baltimore D.;  
RT "Heavy chain variable region contribution to the Npb family of  
antibodies: somatic mutation evident in a gamma 2a variable region.";  
RL Cell 24:625-637(1981).  
CC -1- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY  
RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; J00536; AAA38605.1;  
PIR: A02031; HVMS3.  
DR HSSP; P01810; 2FBJ.  
DR MGD; MGI:96486; Igh-VJ558.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS50835; IG-LIKE; 1.  
KW Immunoglobulin V region; Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 3.  
FT DOMAIN 20 49 FRAMEWORK-1.  
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 55 68 FRAMEWORK-2.  
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 86 117 FRAMEWORK-3.  
FT DISULFID 41 115 BY SIMILARITY.  
FT NON\_TER 117 117  
SQ SEQUENCE 117 AA; 13016 MW; 427C861C53975EDC CRC64;

FT CHAIN 20 117 IG HEAVY CHAIN V REGION 23.  
FT DOMAIN 20 49 FRAMEWORK-1.  
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 55 68 FRAMEWORK-2.  
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 86 117 FRAMEWORK-3.  
FT DISULFID 41 115 BY SIMILARITY.  
FT NON\_TER 117 117  
SQ SEQUENCE 117 AA; 12772 MW; C530F829C906F69B CRC64;  
Query Match 7.2%; Score 11; DB 1; Length 117;  
Best Local Similarity 100.0%; Pred. No. 0.00058;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 60 PGQGLEWIGNI 70  
DB 60 PGQGLEWIGNI 70  
RESULT 9  
ID HV05\_MOUSE STANDARD; PRT; 117 AA.  
AC P01749;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ig heavy chain V region 3 precursor.  
DE IGH-VJ558.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6;  
RX MEDLINE-81234548; PubMed-6788376;  
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,  
RA Baltimore D.;  
RT "Heavy chain variable region contribution to the Npb family of  
antibodies: somatic mutation evident in a gamma 2a variable region.";  
RL Cell 24:625-637(1981).  
CC -1- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY  
RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; J00536; AAA38605.1;  
PIR: A02031; HVMS3.  
DR HSSP; P01810; 2FBJ.  
DR MGD; MGI:96486; Igh-VJ558.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS50835; IG-LIKE; 1.  
KW Immunoglobulin V region; Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 3.  
FT DOMAIN 20 49 FRAMEWORK-1.  
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 55 68 FRAMEWORK-2.  
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 86 117 FRAMEWORK-3.  
FT DISULFID 41 115 BY SIMILARITY.  
FT NON\_TER 117 117  
SQ SEQUENCE 117 AA; 13016 MW; 427C861C53975EDC CRC64;

```
Query Match          7.2%; Score 11; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 0.00058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 PGQGLEWIGNI 70
DB 60 PGQGLEWIGNI 70
|||||

RESULT 10
HV06_MOUSE STANDARD; PRT; 117 AA.
AC P01750;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 102 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPB family of
antibodies; somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -1- MISCELLANEOUS: THIS GERMINE GENE BELONGS TO A SET OF CLOSELY
RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00533; AAA38602.1; -
DR PIR: C90809; HVMS45.
DR HSP: P01810; 2FBJ.
DR MGD; MGI:96486; Igh-VJ558.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 102.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12867 MW; 740A65DD851FCA8C CRC64;

Query Match          7.2%; Score 11; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 0.00058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 SSLTSEDSAVY 113
DB 103 SSLTSEDSAVY 113
|||||

RESULT 11
HV10_MOUSE STANDARD; PRT; 117 AA.
AC P01754; P11270;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ig heavy chain V region 145 precursor.
GN IGH-VJ558.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC MEDLINE=81245215; PubMed=6789211;
RA Givol D., Zakut R., Effron K., Rechavi G., Ram D., Cohen J.B.;
RT "Diversity of germ-line immunoglobulin VH genes.";
RL Nature 292:426-430(1981).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; J00488; AAA38519.1; -  
 DR PIR; A02041; HVMS8A.  
 DR HSSP; P01810; 2FEJ.

DR MGD; MGI:96486; Igh-VJ558.  
 DR InterPro; IPR007110; Ig-Like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 1.

DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS0835; IG-LIKE; 1.  
 KW Immunoglobulin V region; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 117 IG HEAVY CHAIN V REGION 108A.  
 FT DOMAIN 20 >117 IG-LIKE.  
 FT NON\_TER 117 117  
 SQ SEQUENCE 117 AA; 12972 MW; 428CB44DF25D1BC2 CRC64;

Query Match 7.2%; Score 11; DB 1; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 0.00058;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 SSLTSEDSAVY 113  
 |||||

Db 103 SSLTSEDSAVY 113

# RESULT 13

ID HV49\_MOUSE STANDARD; PRT; 117 AA.

AC P06328;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig heavy chain V region VH558 B4 precursor.

OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=8509340; PubMed=2578321;

RA Yancopoulos G.D., Alt F.W.;

RT 'Developmentally controlled and tissue-specific expression of

unrearranged VH gene segments.';

RL Cell 40:271-281(1985).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M13788; AAA38506.1; -

DR PIR; A02035; MHMSB4.

DR HSSP; P01810; 2FEJ.

DR InterPro; IPR007110; Ig-Like.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR003596; Ig\_v.

DR Pfam; PF00047; Ig; 1.

DR SMART; SM00406; IGV; 1.

DR PROSITE; PS0835; IG-LIKE; 1.

KW Immunoglobulin V region; Signal.  
 FT SIGNAL 1 19

FT CHAIN 20 117 IG HEAVY CHAIN V REGION VH558 B4.

FT DOMAIN 20 49 FRAMEWORK-1.

FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.

FT DOMAIN 55 68 FRAMEWORK-2.

FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.

FT DOMAIN 86 117 FRAMEWORK-3.

FT DISULFID 41 115 BY SIMILARITY.

FT NON\_TER 117 117

SQ SEQUENCE 117 AA; 12834 MW; B8862FAC67ABD345 CRC64;

Query Match 7.2%; Score 11; DB 1; Length 117;

Best Local Similarity 100.0%; Pred. No. 0.00058;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 SSLTSEDSAVY 113  
 |||||

Db 103 SSLTSEDSAVY 113

# RESULT 14

ID HV01\_MOUSE

AC P01745;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Ig heavy chain V region MPC 11.

OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=81053741; PubMed=6253904;

RA Zakut R., Cohen J., Givol D.;

RT "Cloning and sequence of the cDNA corresponding to the variable

region of immunoglobulin heavy chain MPC11.1";

RL Nucleic Acids Res. 8:3591-3601(1980).

RN [2]

RP REVISIONS.

RA Zakut R., Cohen J., Givol D.;

RL Nucleic Acids Res. 8:4839-4840(1980).

CC -1- MISCELLANEOUS: THIS SEQUENCE WAS TRANSLATED FROM AN MRNA ISOLATED

CC FROM A MYELOMA THAT SECRETES IGG2B.

CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.

DR PIR; A93708; GVMS11.

DR HSSP; P01810; 2FEJ.

DR InterPro; IPR007110; Ig-Like.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR003596; Ig\_v.

DR Pfam; PF00047; Ig; 1.

DR SMART; SM00406; IGV; 1.

DR PROSITE; PS0835; IG-LIKE; 1.

KW Immunoglobulin V region.

FT DOMAIN 1 112 IG-LIKE.

FT NON\_TER 121 121

SQ SEQUENCE 121 AA; 13135 MW; 227AEF3EC56ED0BF CRC64;

Query Match 7.2%; Score 11; DB 1; Length 121;

Best Local Similarity 100.0%; Pred. No. 0.0006;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LTADTSSSTAY 99  
 -|||||

Db 70 LTADTSSSTAY 80

# RESULT 15

ID HV11\_MOUSE

AC P01755;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUL-1999 (Rel..38, Last annotation update)  
DE Ig heavy chain V region S43 precursor.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=81234548; PubMed=6788376;  
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,  
RA Baltimore D.;  
FT "Heavy chain variable region contribution to the NPb family of  
RT antibodies: somatic mutation evident in a gamma 2a variable region.";  
RL Cell 24:625-637(1981).  
CC -1- MISCELLANEOUS: THE GAMMA-2A CHAIN MRNA WAS CLONED FROM A HYBRIDOMA  
CC MAKING ANTIBODIES TO THE HAPTEN (4-HYDROXY-3-NITROPHENYL)ACETYL  
CC (NPB ANTIBODIES).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J00539; AAA38172.1; -.  
DR PIR; A02038; G2MS43.  
DR HSP; P01810; 2FBJ  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_V.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS50835; IG-LIKE; 1.  
KW Immunoglobulin V region; Signal.  
FT SIGNAL 1  
FT CHAIN 19  
FT DOMAIN 20 137 IG HEAVY CHAIN V REGION S43.  
FT DOMAIN 20 49 FRAMEWORK-1.  
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 55 68 FRAMEWORK-2.  
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 86 117 FRAMEWORK-3.  
FT DOMAIN 118 122 D SEGMENT.  
FT DOMAIN 123 137 JH2 SEGMENT.  
FT DISULFID 41 115 BY SIMILARITY.  
FT NON TER 137 137  
SQ SEQUENCE 137 AA; 15200 MW; ADD5881BF44B8EC9 CRC64;  
  
Query Match 7.2%; Score 11; DB 1; Length 137;  
Best Local Similarity 100.0%; Pred. No. 0.00067;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 103 SSLSSEDSAVY 113  
DB 103 SSLSSEDSAVY 113  
|||||

Search completed: August 30, 2003, 22:10:22  
Job time : 9.24161 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 22:06:06 ; Search time 23.6174 seconds  
(without alignments)  
1671.732 Million cell updates/sec

Title: US-08-836-455-4  
Perfect score: 153  
Sequence: 1 MECSWFLFLSITGVHSQ.....TVSSAKTTPPVYPLVPGSL 153

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- SPTREMBL\_23.\*
- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	10.5	481	11	Q91WT3 mus musculus
2	15	9.8	111	11	Q9DB88 mus musculus
3	15	9.8	426	11	Q9DCD9 mus musculus
4	15	9.8	463	11	Q9DLC4 mus musculus
5	15	9.8	473	11	Q9DBL4 mus musculus
6	14	9.2	109	11	Q9JL75 mus musculus
7	14	9.2	117	11	Q9QXE9 mus musculus
8	14	9.2	117	11	Q9QXF0 mus musculus
9	14	9.2	143	11	Q92AQ0 mus musculus
10	14	9.2	147	11	Q923S3 mus musculus
11	14	9.2	474	11	Q9R3H6 mus musculus
12	13	8.5	110	11	Q9JL77 mus musculus
13	13	8.5	121	11	Q99NG4 mus musculus
14	13	8.5	140	11	Q92AR2 mus musculus
15	13	8.5	142	11	Q924Q1 mus musculus
16	13	8.5	142	11	Q924Q2 mus musculus

17	13	8.5	143	11	Q924P6	Q924p6 mus musculus
18	13	8.5	143	11	Q924R0	Q924r0 mus musculus
19	13	8.5	145	11	Q924Q6	Q924q6 mus musculus
20	13	8.5	145	11	Q924Q9	Q924q9 mus musculus
21	13	8.5	145	11	Q924P7	Q924p7 mus musculus
22	13	8.5	145	11	Q924R1	Q924r1 mus musculus
23	13	8.5	145	11	Q924R4	Q924r4 mus musculus
24	13	8.5	146	11	Q924Q3	Q924q3 mus musculus
25	13	8.5	481	11	Q91WT1	Q91wt1 mus musculus
26	13	8.5	482	11	Q91X92	Q91x92 mus musculus
27	13	8.5	482	11	Q8K172	Q8k172 mus musculus
28	13	8.5	484	11	Q99LA6	Q99la6 mus musculus
29	13	8.5	488	11	Q8K0F2	Q8k0f2 mus musculus
30	13	8.5	489	11	Q8VCX4	Q8vcx4 mus musculus
31	11	7.2	137	11	Q924R6	Q924r6 mus musculus
32	11	7.2	139	11	Q924R5	Q924r5 mus musculus
33	11	7.2	140	11	Q924P8	Q924p8 mus musculus
34	11	7.2	141	11	Q924Q4	Q924q4 mus musculus
35	11	7.2	143	11	Q91VA2	Q91va2 mus musculus
36	11	7.2	143	11	Q924Q5	Q924q5 mus musculus
37	11	7.2	143	11	Q91V67	Q91v67 mus musculus
38	11	7.2	143	11	Q924R7	Q924r7 mus musculus
39	11	7.2	143	11	Q924P9	Q924p9 mus musculus
40	11	7.2	144	11	Q924P5	Q924p5 mus musculus
41	11	7.2	145	11	Q924R3	Q924r3 mus musculus
42	11	7.2	145	11	Q924Q7	Q924q7 mus musculus
43	11	7.2	146	11	Q924R8	Q924r8 mus musculus
44	11	7.2	146	11	Q924Q8	Q924q8 mus musculus
45	11	7.2	168	11	Q8VDC9	Q8vdc9 mus musculus

ALIGNMENTS

RESULT 1

Q91WT3 Q91WT3 PRELIMINARY; PRT; 481 AA.  
AC Q91WT3;  
01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Hypothetical 52.0 kDa protein.  
GN IGH-VJ558 OR AI893585.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Colon;  
RA Strausberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC013488; AAH13488.1; -  
DR MGD; MGI:96486; Igh-VJ558.  
DR InterPro; IPR007110; Igh-Like.  
DR InterPro; IPR003006; Igh\_MHC.  
DR InterPro; IPR003596; Igh\_v.  
DR Pfam; PF00047; Igh; 4.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IGH\_LIKE; 4.  
DR PROSITE; PS00290; IGH\_MHC; 2.  
KW Hypothetical protein.  
SQ SEQUENCE 481 AA; 52022 MW; 4EBB5C253038B718 CRC64;

Query Match 10.5%; Score 16; DB 11; Length 481;  
Best Local Similarity 100.0%; Pred. No. 7.1e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 103 SSSLTSEDSAVYFCARG 118

Db 102 SSSLTSEDSAVYFCARG 117



Qy 103 SSSLTSEDSAVYFCAR 117  
 Db 103 SSSLTSEDSAVYFCAR 117

## RESULT 5

Q9D8L4 PRELIMINARY; PRT; 473 AA.  
 AC Q9D8L4;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE 181006009Rik protein.  
 GN IGH-1 OR 181006009R1K.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK007918; BAB25349.1; -  
 DR HSP; P01842; 7FAB.  
 DR MGD; MGI:96443; Igh-1.  
 DR InterPro; IPR007110; Igh-like.  
 DR InterPro; IPR003006; Igh.MHC.  
 DR InterPro; IPR003596; Igh-v.  
 DR Pfam; PF00047; Igh; 4.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 4.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match 9.8%; Score 15; DB 11; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-07;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 103 SSSLTSEDSAVYFCAR 117  
 Db 103 SSSLTSEDSAVYFCAR 117

## RESULT 6

Q9JL75 PRELIMINARY; PRT; 109 AA.  
 AC Q9JL75;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Anti-myosin immunoglobulin heavy chain variable region  
 DE (Fragment).  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RX MEDLINE=20448942; PubMed=10992488;  
 RA Malkiel S., Liao L., Cunningham M.W., Diamond B.;  
 RT "T-Cell-dependent antibody response to the dominant epitope of  
 streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive  
 with cardiac myosin.";  
 RL Infect. Immun. 68:5803-5808(2000).  
 DR EMBL; AF206031; AAF69329.1; -  
 DR HSP; P01810; 2FBJ.  
 DR InterPro; IPR007110; Iq-like.  
 DR InterPro; IPR003006; Iq.MHC.  
 DR InterPro; IPR003596; Iq-v.  
 DR Pfam; PF00047; Iq; 1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR NON\_TER 1  
 FT NON\_TER 109 109  
 SQ SEQUENCE 109 AA; 12118 MW; FF65E441BBF936A6 CRC64;

Query Match 9.2%; Score 14; DB 11; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 GASVKMSCKASGYT 47  
 Db 6 GASVKMSCKASGYT 19

## RESULT 7

Q9QXE9 PRELIMINARY; PRT; 117 AA.  
 AC Q9QXE9;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Immunoglobulin heavy chain V-D-J region (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Clemens A., Rademaekers A., Specht C., Koelsch E.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ225174; CAB65237.1; -  
 DR HSP; P01810; 2FBJ.  
 DR InterPro; IPR007110; Iq-like.  
 DR InterPro; IPR003006; Iq.MHC.  
 DR InterPro; IPR003596; Iq-v.  
 DR Pfam; PF00047; Iq; 1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR NON\_TER 1  
 FT NON\_TER 117 117  
 SQ SEQUENCE 117 AA; 13000 MW; CDDE2AF84D499734 CRC64;

Query Match 9.2%; Score 14; DB 11; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 GASVKMSCKASGYT 47  
 Db 15 GASVKMSCKASGYT 28

## RESULT 8

Q9QXF0 PRELIMINARY; PRT; 117 AA.  
 ID Q9QXF0

```

AC Q9QXF0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Immunoglobulin heavy chain V-D-J region (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Clemens A., Rademakers A., Specht C., Koelsch E.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ225171; CAB65236.1; -
DR HSP; P01789; IMCP.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 13060 MW; D816AD0858A47E4C CRC64;

Query Match          9.2%; Score 14; DB 11; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 GASVKMSCKASGYT 47
DB 15 GASVKMSCKASGYT 28

RESULT 9
Q924Q0
ID Q924Q0 PRELIMINARY; PRT; 143 AA.
AC Q924Q0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE V165-D-J-C mu protein (Fragment).
GN V165-D-J-C MU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kozono Y., Kozono H., Azuma T.;
RL "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB069915; BAB63931.1; -
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 143
SQ SEQUENCE 143 AA; 15704 MW; C99D2433F2BAD8A0 CRC64;

Query Match          9.2%; Score 14; DB 11; Length 143;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 GASVKMSCKASGYT 47
DB 15 GASVKMSCKASGYT 28

Q925S3
ID Q925S3 PRELIMINARY; PRT; 147 AA.
AC Q925S3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MRP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,
RL "Mechanism of exogenous nucleic acids and their precursors improving
RT the repair of intestinal epithelium after irradiation in mice.";
RL World J. Gastroenterol. 6:709-717(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-BALB/c;
RA Cui D., Zeng G., Yan X., Li X., Su C.;
RL "Cloning of mouse genes related to repairing of intestinal epithelium
RT of the irradiated mice by treatment with the intestinal RNA of mice.";
RT the same 'strain.';
RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).
DR EMBL; AF240166; AAK43731.1; -
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
SQ SEQUENCE 147 AA; 16274 MW; 800594A12B97191F CRC64;

Query Match          9.2%; Score 14; DB 11; Length 147;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LTSSEDSAVYFCARG 118
DB 88 LTSSEDSAVYFCARG 101

RESULT 11
Q8R3H6
ID Q8R3H6 PRELIMINARY; PRT; 474 AA.
AC Q8R3H6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 51.7 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025447; AAK25447.1; -
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.

```

DR PROSITE; PS50835; IG\_LIKE; 4.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 474 AA; 51748 MW; 8608B57C6CD2874A CRC64;

Query Match 9.2%; Score 14; DB 11; Length 474;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 SLTSEDSAVYFCAR 117  
 |||||  
 DB 104 SLTSEDSAVYFCAR 117

RESULT 12  
 Q9JL77 PRELIMINARY; PRT; 110 AA.

ID AC Q9JL77;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Anti-myosin immunoglobulin heavy chain variable region  
 (Fragment).  
 DE DE  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DBA/2;  
 RX MEDLINE=20448942; PubMed=10922488;  
 RA Maikael S., Liao L., Cunningham M.W., Diamond B.;  
 RT "a-Cell-dependent antibody response to the dominant epitope of  
 RT streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive  
 RT with cardiac myosin".  
 RL Infect. Immun. 68:5803-5808(2000).  
 DR EMBL; AF206029; AAF69327.1; -;  
 DR HSP; P01810; 2FBJ.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF000047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS500835; IG\_LIKE; 1.  
 FT NON\_TER 1  
 FT NON\_TER 110  
 SQ SEQUENCE 110 AA; 12138 MW; 2EDE81FB5862C9AF CRC64;

Query Match 8.5%; Score 13; DB 11; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 DYWGQGTSTVTSS 137  
 |||||  
 DB 98 DYWGQGTSTVTSS 110

RESULT 13  
 Q99NG4 PRELIMINARY; PRT; 121 AA.

ID AC Q99NG4;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Single chain Fv (Fragment).  
 DE OS  
 OS Mus musculus (Mouse).  
 OG Plasmid pHEN1.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALE/C;

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE V23-D-J-C mu protein (Fragment).  
 GN V23-D-J-C MU.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RA Kozono Y., Kozono H., Azuma T.;  
 RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals  
 RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-  
 RL Hydroxy-3-Nitrophenyl)Acetyl (NP).";  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB069913; BAB63929.1;  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 DR PROSITE: PS50835; IG\_LIKE; 1.  
 FT NON\_TER 1  
 FT NON\_TER 142  
 FT SEQUENCE 142 AA; 15622 MW; 24A265CE4EA4318B CRC64;  
 SQ SEQUENCE 142 AA; 15622 MW; 24A265CE4EA4318B CRC64;

Query Match 8.5%; Score 13; DB 11; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 125 DYWGQGTSTVTVSS 137  
 Db 105 DYWGQGTSTVTVSS 117  
 |||||

Search completed: August 30, 2003, 22:11:19  
 Job time : 24.6174 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 30, 2003, 22:10:26 ; Search time 3130.85 Seconds  
(without alignments)  
1999.191 Million cell updates/sec

Title: US-08-836-455-4  
Perfect score: 816  
Sequence: 1 MECNVFLFLSITGVHSQ.....TVSSAKTTPPVYPLVPGSL 153

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus.p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO.spool/US08836455/runat\_29082003\_132901\_22281/app\_query.fasta\_1.654  
-DB=GenEmbl -QFWT-fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdl -LIST=45  
-DOALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US08836455\_@CGN\_1\_1\_3608\_@runat\_29082003\_132901\_22281 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=7  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl :  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pin.\*  
35: em.htg\_rod.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em.sy.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	816	100.0	461	6	AR164506 Sequence
2	816	100.0	461	6	BD085738 Methods o
3	725	88.8	411	10	AF124720 Mus muscu
4	630	77.2	471	10	AB050080 Mus muscu
5	621	76.1	1553	6	E08434 cDNA encodi
6	609	74.6	483	10	U39899 Mus musculu
7	603	73.9	1542	10	BC002121 Mus muscu
8	602	73.8	525	10	M28834 Mus musculu
9	602	73.8	540	6	A13735 variable re
10	601.5	73.7	1526	10	MMU555622 Mus muscu
11	599.5	73.5	1544	10	MUSIGBLH1
12	597.5	73.2	457	10	AY178830 Mus muscu
13	595	72.9	489	10	MMU39900 Mus muscu
14	590	72.3	489	10	AB050070 Mus muscu
15	581	71.2	1557	10	BC003435 Mus muscu
16	580	71.1	473	10	AB050071 Mus muscu
17	575	70.5	1573	10	BC018280 Mus muscu
18	574.5	70.4	1568	6	E33134 Humanized a
19	574.5	70.4	1570	6	A22261 M.musculus
20	574.5	70.4	1570	6	A77138 Sequence 6
21	569	69.7	1620	10	BC003888 Mus muscu
22	568.5	69.7	453	6	AR109947 Sequence
23	568.5	69.7	509	10	M1903 Mouse Ig re
24	568	69.6	1579	10	X13188 Mouse mRNA
25	568	69.6	1683	6	E35543 Transgenic
26	567.5	69.5	416	10	AF045892 Mus muscu
27	567.5	69.5	902	12	U49832 Synthetic s
28	566.5	69.4	450	10	AY178829 Mus muscu
29	565.5	69.3	1570	6	AR029102 Sequence
30	565	69.2	458	6	I05921 Sequence 37
31	565	69.2	458	6	I08811 Sequence 12
32	565	69.2	458	6	I09199 Sequence 38
33	565	69.2	458	10	M17953 Mouse Ig re
34	563	69.0	1500	10	M79907 M.musculus
35	563	69.0	1581	6	A78881 Sequence 1
36	563	69.0	1581	10	X70423 M.musculus
37	563	69.0	1629	10	BC033451 Mus muscu
38	562.5	68.9	460	10	AB050073 Mus muscu
39	562	68.9	405	10	AB069863 Mus muscu
40	561.5	68.8	1083	6	AX556950 Sequence
41	561.5	68.8	1083	6	AX709549 Sequence
42	561.5	68.8	1413	6	AX556949 Sequence
43	561.5	68.8	1413	6	AX709548 Sequence
44	561.5	68.8	1558	10	AF466769 Mus muscu
45	561.5	68.8	9209	6	AR000007 Sequence

ALIGNMENTS

RESULT 1

AR164506	AR164506	Sequence 3 from patent US 6274143.	461 bp	DNA	linear	PAT 17-OCT-2001
LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	
UNKNOWN.	UNKNOWN.	UNCLASSIFIED.	1 (bases 1 to 461)	Chatterjee,M. and Poon,K.A. Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11D10	Patent: US 6274143-A 3 14-AUG-2001; Location/Qualifiers 1..461 /organism="unknown"	
BASE COUNT	112 a	121 c	119 g	109 t		
ORIGIN						
Alignment Scores:						
Pred. No.:	5.55e-75	Length:	461			
Score:	816.00	Matches:	153			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	100.00%	Indels:	0			
DB:	6	Gaps:	0			
US-08-836-455-4 (1-153) x AR164506 (1-461)						
Qy	1 MetGlcCysSerTrpValPheLeuPheLeuSerIleThrThrGlyValHisSerGln	20				
Db	1 ATGAATGCAGCTGGGTCTTCTCTCTCTCATTAATACTAGAGTGTCACCTCCAG	60				
Qy	21 AlaTyrrLeuGlnInSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer	40				
Db	61 GCTATTCTACAGCAGCTCGGGCTGAGCTGGTAGGTTGAGGCTGGGGCTCAGTGAAGATGCC	120				
Qy	41 CysLysAlaSerGlyTyrrLeuThrSerTyrrAsnMethIstTpValLysGlnThrPro	60				
Db	121 TGCAGGCTTCTGGCTACACATTGACCAGTTACAAATATGCACCTGGGTAAAGCACACACT	180				
Qy	61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrrAsn	80				
Db	181 GGACAGGGCTGGAATGGATTGGAAATATTTTTCTCGAAATGCTACTTACTACAAT	240				
Qy	81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrrMet	100				
Db	241 CAGAAGTTTAAGGCAAGGCTCATTCACATGCAGACACATCCTCCAGCAGACCTACATG	300				
Qy	101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrrPheCysAlaArqGlyAsnTrp	120				
Db	301 CAGATCAGCAGCCTGACATCTGAAGACTCGCGTCTATTCTGTGCAAGAGGAACCTGG	360				
Qy	121 GluGlyAlaLeuAspTyrrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThr	140				
Db	361 GAGGGTGCTCTGGACTCTGGGCTCAAGAACCTCAGTCCCGTCTCTCTCAGCCAAAACG	420				
Qy	141 ThrProProValTyrrProLeuValProGlySerLeu	153				
Db	421 ACACCCCCCCTCTATCCACTGGTCCCTGGGAAGCTTG	459				
RESULT 2	BD085738					
LOCUS	DEFINITION	461 bp	DNA	linear	PAT 27-AUG-2002	
LOCUS	Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11D10.					
ACCESSION	BD085738					
VERSION	BD085738.1	GI:22631348				
KEYWORDS	JP 2001523269-A/2.					
SOURCE	unidentified					
ORGANISM	unclassified.					



LOCUS AF124720 411 bp mRNA linear ROD 22-MAY-2001  
DEFINITION Mus musculus immunoglobulin heavy chain mRNA, partial cds.  
ACCESSION AF124720  
VERSION AF124720.1 GI:14164544  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Tripathi, P.K., Qin, H., Bhattacharya-Chatterjee, M., Ceriani, R.L.,  
Foon, K.A. and Chatterjee, S.K.  
TITLE Construction and characterization of a chimeric fusion protein  
consisting of an anti-idiotypic antibody mimicking a breast  
cancer-associated antigen and the cytokine GM-CSF  
JOURNAL Hybridoma 18 (2), 193-202 (1999)  
MEDLINE 99306687  
PUBMED 10380019  
REFERENCE 2 (bases 1 to 411)  
AUTHORS Chatterjee, S.K. and Tripathi, P.K.  
TITLE Direct Submission  
JOURNAL Submitted (29-JAN-1999) Internal Medicine, University of Kentucky,  
800 Rose Street, Lexington, KY 40536, USA  
FEATURES  
source  
1..411  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
1..>411  
/note="anti-idiotypic antibody 11D10; mimics a breast  
cancer-associated antigen, human fat globule (HMFg)"  
/codon\_start=1  
/product="immunoglobulin heavy chain"  
/protein\_id="AAK55119.1"  
/db\_xref="GI:14164545"  
/translation="MKCSWVFLFLSITGVHSQAYLQSGAELVRSAGSVKMSCKAS  
GYTLTSNHHVYKQTGGQLEWIGYIFLNAGTNYNOKFKGASLTADTSSSTAYMQI  
SSLTSDSAVFCARGNWEGLDYGQGTSTVSS"  
BASE COUNT 102 a 101 c 107 g 101 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.12e-65 Length: 411  
Score: 725.00 Matches: 136  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.27% Mismatches: 0  
Query Match: 88.85% Indels: 0  
DB: 10 Gaps: 0  
US-08-836-455-4 (1-153) x AF124720 (1-411)  
QY 1 MetGluCysSerTrpValPheLeuSerIleThrThrGlyValHisSerGln 20  
Db 1 ATGAATGACGTGGGTCTTCCTCCCTGCTCAATACTACAGGTGCCACTCCGAC 60  
QY 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
Db 61 GCTTATCTACAGCAGCTGGGGCTGAGCTGGTGGCTGGGCGCTCAGTAGAGTGTC 120  
QY 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstTrpValLysGlnThrPro 60  
Db 121 TGAAGGCTTCTGGCTACACATTGACAGTACATATGACATGACATGACATGACACCT 180  
QY 61 GlyGlnGlyLeuGlnTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrAsn 80  
Db 181 GGACAGGGCTGGAATGGATGGAAATATTTTCCTGGAAATGGTGATCTACTACAAT 240  
QY 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerThrAlaTyrMet 100  
Db 241 CAGAAGTTTAAGCGAGCGCTCATTTGACTGACGACACATCTCCACGACAGCCTACATG 300  
QY 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120

Db 301 CAGATCAGCAGCCTGACATCTGAAGACTCTCGGCTCTATTCTCTGTCAGAGGGAAGTGG 360  
QY 121 GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSer 137  
Db 361 GAGGTGCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCA 411  
RESULT 4  
LOCUS AB050080 471 bp mRNA linear ROD 02-APR-2002  
DEFINITION Mus musculus VH9H8 mRNA for anti-A/U antibody, partial cds.  
ACCESSION AB050080  
VERSION AB050080.1 GI:19909935  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Kitagawa, Y. and Okuhara, E.  
TITLE The separation of three antibody populations from  
anti-poly(A).poly(U) antibodies elicited in mice or rabbits and  
antigenic features of poly(A).poly(U)  
JOURNAL Mol. Immunol. 19 (2), 257-266 (1982)  
MEDLINE 82245325  
PUBMED 6178956  
REFERENCE 2  
AUTHORS Kitagawa, Y.  
TITLE anti-dsRNA (A/U) Ab VH region VH9H8  
JOURNAL Published Only in Database (2002)  
REFERENCE 3 (bases 1 to 471)  
AUTHORS Kitagawa, Y.  
TITLE Direct Submission  
JOURNAL Submitted (17-OCT-2000) Yoshichika Kitagawa, Akita Prefectural  
University, Biotechnology Institute; Minami 2-2, Ogata, Akita  
010-0444, Japan (E-mail: kitagawa@agri.akita-pu.ac.jp,  
URL: www.akita-pu.ac.jp/, Tel: 81-185-45-2026 (ex.400),  
Fax: 81-185-45-2678)  
FEATURES  
source  
1..471  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
1..471  
/cell\_line="hybridoma 9H8"  
/gene="VH9H8"  
11..>469  
/gene="VH9H8"  
/codon\_start=1  
/product="anti-A/U antibody"  
/protein\_id="BAB87192.1"  
/db\_xref="GI:19909936"  
/translation="MGWTGVFLILLITGVHSQAYLQSGAELVRSAGSVKMSCKAS  
GYTLTSNHHVYKQTGGQLEWIGYIFLNAGTNYNOKFKGASLTADTSSSTAYMQIS  
SLTSDSAVFCARGNWEGLDYGQGTSTVSSAARTTAQVYPLAPGSL"  
BASE COUNT 114 a 126 c 121 g 110 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 7.75e-56 Length: 471  
Score: 630.00 Matches: 128  
Percent Similarity: 85.26% Conservative: 5  
Best Local Similarity: 82.05% Mismatches: 19  
Query Match: 77.21% Indels: 4  
DB: 10 Gaps: 3  
US-08-836-455-4 (1-153) x AB050080 (1-471)  
QY 1 MetGluCysSerTrp-----ValPheLeuPheLeuSerIleThrThrGlyValHis 18  
Db 5 GTCGACATGGGATGGAGCCGGGGTCTTCCTCATCTCTGCAATAACTACAGGTGTCAC 64

```

Qy 19 SerGlnAlaTyrIleuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLys 38
Db 65 TCCAGGCTTATCTACACAGCTCTGGGGCTGAACCTGGTGGCTGGGGCTCAGTGAAG 124

Qy 39 MetSerCysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMetHisTrrpValLysGln 58
Db 125 ATGTCCTGCAAGGCTCTGGCTACACATTACCAAGTTACAATATGCACTGGGTAAAGCAG 184

Qy 59 ThrProGlyGlnGlyLeuGluTrrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyr 78
Db 185 ACACCTGGACAGGCGCTGGAATGATGGATATATTTTCCTG---AATGCTGGTACTAAC 241

Qy 79 TyrAsnGlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerThrAla 98
Db 242 TACAATCAGAAGTTCAAGGGCAAGGCCACATTTGACCGCAGACACATCTCCAGCACAGCC 301

Qy 99 TyrMetGlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGly 118
Db 302 TACATGCAGATCAGCAGCCTGACATCTGAAGACTCTCGGTCTATTCTGTGCAAGAGGG 361

Qy 119 AsnTrrpGluGlyAla---LeuAspTyrTrrpGlyGlnGlyThrSerValThrValSerSer 137
Db 362 GAGTATGGTAAACCCCTGGTTGCTTACTGGGGCCAGGGACTCTGGTCACTGCTCTGCA 421

Qy 138 AlaLysThrThrProProValTrrpProLeuValProGlySerLeu 153
Db 422 GCCAAACACAGCCCAACCGTTTATCCCTTGGCCCTGGAGCTTG 469

RESULT 5
E08434
LOCUS E08434 1553 bp DNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding H-chain of anti-TMV-antibody.
ACCESSION E08434
VERSION E08434.1 GI:2176551
KEYWORDS JP 1994319396-A/7.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1553)
AUTHORS Saito,Y., Kogiku,T., Kamishiro,T., Murafuji,H., Takami,M. and Fumino,M.
TITLE PLANT FOR PRODUCTION ANTIVIRAL ANTIBODY AND METHOD FOR CREATING THE SAME PLANT
JOURNAL Patent: JP 1994319396-A 7 22-NOV-1994; JAPAN TOBACCO INC, KURARAY CO LTD
COMMENT OS None
OC Artificial sequences.
PN JP 1994319396-A/7
PD 22-NOV-1994
PF 07-MAY-1993 JP 1993131208
PI SAITO YASUO, KOGIKU TOSHIHIKO, KAMISHIRO TAKASHI, PI MURAFUJI HITOAKI, TAKAMI MASAAKI, FUMINO MASAYASU
PC A01H5/00,C12N15/13,C12P21/08;
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..1553
FT /organism='Artificial sequences' FT
misc_feature 1..1553
FT anti-TMV-antibody, /note='cDNA encoding H-chain of
FT 5'UTR 1..55
FT CDS 56..1453
FT /product='H-chain of anti-TMV-antibody' FT
3'UTR 1454..1553.
Location/Qualifiers
1..1553
/organism='unidentified'
/mol_type='genomic DNA'
FEATURES
source

```

```

BASE COUNT 386 a 450 c 372 g 345 t
ORIGIN
Alignment Scores:
Pred. No.: 2,52e-54 Length: 1553
Score: 621.00 Matches: 118
Percent Similarity: 85.26% Conservative: 15
Best Local Similarity: 75.64% Mismatches: 19
Query Match: 76.10% Indels: 4
DB: 6 Gaps: 2

US-08-836-455-4 (1-153) x E08434 (1-1553)
Qy 1 MetGluCysSerTrrpValPheLeuPheLeuLeuSerIleThrThrGlyValHisSerGln 20
Db 56 ATGAATGTAACGGATACCTTCCTTTTATCTGTCTCACTAACTTCAGGTCTACTCAG 115
Qy 21 AlaTyrLeuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 116 GTTCAGCTCCAGCAGCTCTGGGGCTGAGCTGGCAAGACCTGGGCTTCAGTGAAGTTGCC 175
Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMetHisTrrpValLysGlnThrPro 60
Db 176 TGAAGGCTTCTGGCTACACCTTTTACTAGCTAGCTGGATGAGCTGGGTAAACACAGGCTC 235
Qy 61 GlyGlnGlyLeuGluTrrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80
Db 236 GGACAGGCTCTGGAATGATGGGGCTATTATCTCTGGAATGCTGATAGTACTACT 295
Qy 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100
Db 296 CAGAAGTTCAAGGGCAAGGCCACATTTACTGCTGACATAAATCTCCACAGCAGCCTACATG 355
Qy 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGly----- 118
Db 356 CAACTCAGCGCTTGGCATCTGAGGACTCTGCGGTCTATTACTGTGCAAGAGAGGGGGT 415
Qy 119 ---AsnTrrp---GluGlyAlaLeuAspTrrpGlyGlnGlyThrSerValThrValSer 136
Db 416 TACTCTGGTCCGACTATGCTATGAGCTACTGGGTCAAGGAACCTCAGTCACGCTCTCC 475
Qy 137 SerAlaLysThrThrProProValTrrpProLeuValProGlySer 152
Db 476 TCAGCCAAACACACACCCCATCTGTCTATCCACTGGCCCTGGATCT 523

RESULT 6
MMU39899 483 bp DNA linear ROD 21-NOV-1996
LOCUS MMU39899
DEFINITION Mus musculus anti-glycoprotein-B of human Cytomegalovirus immunoglobulin Vh chain gene, partial cds.
ACCESSION U39899
VERSION U39899.1 GI:1680666
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 483)
AUTHORS Schoppel,K., Hassfurth,E., Britt,W., Ohlin,M., Borrebaeck,C.A. and Mach,M.
TITLE Antibodies specific for the antigenic domain 1 of glycoprotein B (gpUL55) of human cytomegalovirus bind to different substructures
JOURNAL Virology 216 (1), 133-145 (1996)
MEDLINE 96187797
REFERENCE 8614980
AUTHORS 2 (bases 1 to 483)
Schoppel,K.
Direct Submission
Submitted (02-NOV-1995) Michael Mach, Institute of Virology, University Erlangen-Nuernberg, Schlossgarten 4., 91054 Erlangen, Germany
Location/Qualifiers
FEATURES

```

```
source 1. .483
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/note="monoclonal antibody 27-156"
1. .5483
/codon_start=1
/product="anti-glycoprotein-B of human Cytomegalovirus
immunoglobulin Vh chain"
/protein_id="AAB26952.1"
/db_xref="GI:1680667"
/translation="MEWTVILFLSVTEGVHSQVLOQSGAELVRPGSSVKISKAS
GYAFSGYINWVYKPGQGLEWYIGQYIPGDGTDYNGFKRATITADKSSSTAYMQL
SSLTSEDSAYVEFCARGYDEVYFYFADYWGQTSVTSVSAKTTPTSPVPLAPVCGDKL
G"
BASE COUNT 117 a 118 c 130 g 118 t
ORIGIN

Alignment Scores:
Pred. No.: 1.15e-53 Length: 483
Score: 609.00 Matches: 117
Percent Similarity: 85.06% Conservative: 14
Best Local Similarity: 75.97% Mismatches: 19
Query Match: 74.63% Indels: 4
DB: 10 Gaps: 1

US-08-836-455-4 (1-153) x MMU39899 (1-483)
QY 1 MetGluCysSerTrpValPheLeuPheLeuSerIleThrThrGlyValHisSerGln 20
DB 1 ATGGAAATGACCGCTGGGTCTCTTCCTCCCTGTCAGTAACGAGGTGTCACATCCGAG 60
QY 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
DB 61 GTTCAGCTCCAGCAGCTCTGGGCTGAGCTGGTGAGCCCTGGGTCCCTCAGTAGATTTC 120
QY 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstTrpValLysGlnThrPro 60
DB 121 TGCAGAGCTCTGGCTATGCATTCAGTCTAGTCTAGTAACTGGCTGAAGCAGAGGCT 180
QY 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80
DB 181 GGACAGGGCTTTCAGTGGATGGACAGATTTTCCTGGAGATGGTGATGACTGACTACAAT 240
QY 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerThrAlaTyrMet 100
DB 241 GGAAATTCAGGGTAAAGCCACACTGACTGACAGACAATCTCCAGCAGCAGCTACATG 300
QY 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGly----- 118
DB 301 CAGCTCAGCAGCCTAACATCTGAGGACTCTGGGCTCTATTTCGTGCAAGAGGTTACGAC 360
QY 119 -----AsnTrpGluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSer 136
DB 361 GAAGTGGATATTACTTCTGAGGACTACTGGGGTCAAGGAACCTCAGTCCACCGTCTCC 420
QY 137 SerAlaLysThrThrProProValTyrProLeuValPro 150
DB 421 TCAGCCAAACACACACCCCATCAGTCTATCCACTGGCTCCT 462

RESULT 7
BC002121 1542 bp mRNA linear ROD 16-APR-2003
LOCUS Mus musculus immunoglobulin heavy chain 4 (serum IgG1), mRNA (cDNA
DEFINITION BC002121
ACCESSION BC002121
VERSION BC002121.1 GI:12805308
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

REFERENCE  
AUTHORS

1 (bases 1 to 1542)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altshul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Hellon, B., Kettman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REMARK

## COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcdpaxil@stanford.edu](mailto:mcdpaxil@stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 8 Row: a Column: 21  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis, similarity but not identity to protein.

## FEATURES

## source

1. .1542  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CZECH II"  
/db\_xref="taxon:10090"  
/clone="MGC:6628 IMAGE:3491766"  
/tissue\_type="Mammary tumor metastasized to lung. Tumor  
arose spontaneously from a senescent normal mammary  
(clonal) outgrowth infected with the virus MMV."  
/clone\_lib="NCI\_CGAP\_Lu29"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1. .1542  
/genes="Igh-4"  
/note="synonym: IgG1"  
/db\_xref="LocusID:16017"  
/db\_xref="MGI:96446"  
33. .1418  
/codon\_start=1  
/product="Igh-4 protein"  
/protein\_id="AAH02121.1"  
/db\_xref="GI:12805309"

## gene

## CDS

/db\_xref="LocusID:16017"  
 /translation="MEWIFLFLILSGTAVSHSQVLQOQSGAELARPGASVKMSCKAS  
 GYFTSGISWVKQYRGLLEIYPGSNTYNEKFKGKATLADKSSSTAYMOL  
 TSLTSEDVAIFCARMECPWDVYWGOTTITVSSAKTTPSVYPLAPGSAAPTNSV  
 TLGLVKGYPEPTVTWNSGLSSGVTTPAVLQSLDLYTLSSSVTVPSSTWBPQVTV  
 CNAHPASSKFKDKLIVRCGCKPCICTYVESSVFIFPKPKDVLITLTKPVTCV  
 VVDSKDDPEVQESWFDVVEHTAQTKEEQFNSTFRSSELPINHQDMLNGKEPK  
 CRYNSAFPAPIETKISKTGRPAQOVITPPPEQMAKDKVSLTCTMTDFFPEDIT  
 VEWNGQPAENKNTQPIIMDTDGSYEVYSKLVQKSNWEAGNTFTCSVLHREGLNHH  
 TEKLSHSPCK"  
 BASE COUNT 406 a 447 c 364 g 325 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.78e-52 Length: 1542  
 Score: 603.00 Matches: 114  
 Percent Similarity: 83.55% Conservative: 13  
 Best Local Similarity: 75.00% Mismatches: 25  
 Query Match: 73.90% Indels: 0  
 DB: 10 Gaps: 0

US-08-836-455-4 (1-153) x BC002121 (1-1542)

QY 1 MetGlucySerrtrpValPheLeuLeuSerileThrThrGlyValHisSerGln 20  
 Db 33 ATGAATGGATCTGGATCTTCTTCATCTCTCAGAACTGCAAGTGCCACTCCAG 92  
 QY 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
 Db 93 GTTCAGCTGCAGCAGCTGGAGCTGAGCTGGCAGGCGCTGGGCTTCAGTGAAGATGCC 152  
 QY 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethHisTrpValLysGlnThrPro 60  
 Db 153 TGAAGGCTTCTGGTACTATTCACAAAGCTATGTGATAAGCTGGTGAAGCAGCAAACT 212  
 QY 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80  
 Db 213 AGCAGGGCTTGGATGGAGATTTATCTCTGGAAGTGGTAACACTTACTACAAT 272  
 QY 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100  
 Db 273 GAAAGTTCAGGCGCAAGCCACACTGACTGCAGACAAATCCCTCCAGCAGCCTACATG 332  
 QY 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120  
 Db 333 CAGCTCAGCAGCTGACATCTGAGGACTGCGAGTCTATTCTCTGCAAGATGGGGGAG 392  
 QY 121 GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThr 140  
 Db 393 CCTGGGAGCTGACTACTGGGCGCAAGCCACACTATCATCAGTCTCTCTCCAGCAAAAG 452  
 QY 141 ThrProProValTyrProLeuValProGlySer 152  
 Db 453 ACACCCCATCTGTCTATCCACTGGCCCTGGATCT 488

# RESULT 8

MUSIGHALPA  
 LOCUS 525 bp mRNA linear ROD 18-NOV-1994  
 DEFINITION Mus musculus IgG2a chain (anti-pseudomonas aeruginosa lipoprotein I antibody) mRNA, 5' end.  
 ACCESSION M28834  
 VERSION M28834.1 GI:576597  
 KEYWORDS V-segment; anti-lipoprotein antibody; immunoglobulin G2a gamma chain; immunoglobulin heavy chain.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Marget,M., Eckhardt,A., Ehret,W., von Specht,B.U., Duchene,M. and Domdey,H.  
 Cloning and characterization of cDNAs coding for the heavy and light chains of a monoclonal antibody specific for Pseudomonas

aeruginosa outer membrane protein I  
 Gene 74 (2), 335-345 (1988)  
 MEDLINE 89232725  
 PUBMED 3149944  
 COMMENT On Nov 28, 1994 this sequence version replaced gi:342018.  
 Original source text: Mus musculus (strain BALB/c, sub\_species domesticus) hybridoma cDNA to mRNA.  
 FEATURES  
 Location/Qualifiers  
 source 1..525  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/c"  
 /sub\_species="domesticus"  
 /db\_xref="taxon:10090"  
 /cell\_line="Mab 6A4"  
 /tissue\_type="hybridoma"  
 gene 1..525  
 /gene="Igg"  
 CDS 49..525  
 /gene="Igg"  
 /codon\_start=1  
 /product="immunoglobulin gamma-2a chain"  
 /protein\_id="AA53291.1"  
 /db\_xref="GI:576598"  
 /translation="MERHWLFLEFSVTAGVHSQVLQOQSGAELARPGASVKMSCKAS  
 GYFTTAYMHVVKQRPQGLEWIGYINPTGYTEYNQNFKDKATLTADKSSSTAYMOL  
 SLLTSEDVAIYCTRSYINYEAGMDYWGQCTSVTSAAKTTPSVYPLAPVCGDTTG"  
 49..105  
 sig\_peptide /gene="Igg"  
 106..525  
 mat\_peptide /gene="Igg"  
 /product="immunoglobulin gamma-2a chain"  
 BASE COUNT 144 a 142 c 119 g 120 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 6.65e-53 Length: 525  
 Score: 602.00 Matches: 116  
 Percent Similarity: 82.24% Conservative: 9  
 Best Local Similarity: 76.32% Mismatches: 25  
 Query Match: 73.77% Indels: 2  
 DB: 10 Gaps: 1  
 US-08-836-455-4 (1-153) x MUSIGHALPA (1-525)  
 QY 1 MetGlucySerrtrpValPheLeuLeuSerileThrThrGlyValHisSerGln 20  
 Db 49 ATGAAGGCGACTGGATCTTCTCTCTGTTTCACTGCTGAGTGCCACTCCAG 108  
 QY 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
 Db 109 GTCCAGCTTCAGCAGTCTGGGCTGAACTGGCAAAACCTGGGCTCAGTGAAGATGCC 168  
 QY 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethHisTrpValLysGlnThrPro 60  
 Db 169 TGAAGGCTTCTGGCTACACCTTTACTGCTGATGCTGGTGGTAAACAGAGCCCT 228  
 QY 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80  
 Db 229 GGACAGGCTCTGGAATGGATTGGATACATTAACTTAACTGCTGATGATACAT 288  
 QY 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100  
 Db 289 CAGAAGTTCAGGACCAAGGCCACATTGACTGCAGACAAATCCCTCCAGCAGCCTACATG 348  
 QY 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGly----- 118  
 Db 349 CAAGTGAAGCCTGACATCTGAGGACTCTGAGTCTATTATTGTCACAGAAAGCTACTAT 408  
 QY 119 AsnTrpGluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAla 138  
 Db 409 AACTAGAGGGGCTATGGACTACTGGGCTCAGGAACCTCAGTACCCGCTCTCTCAGCC 468



```

polyA_signal 1475..1480
/gene="IGHG1"
polyA_site 1510
/gene="IGHG1"
BASE COUNT 396 a 436 c 360 g 334 t
ORIGIN
Alignment Scores:
Pred. No.: 2.5e-52 Length: 1526
Score: 601.50 Matches: 115
Percent Similarity: 82.58% Conservative: 13
Best Local Similarity: 74.19% Mismatches: 24
Query Match: 73.71% Indels: 3
DB: 10 Gaps: 1

US-08-836-455-4 (1-153) x MMU555622 (1-1526)

Qy 1 MetGluCysSerTrpValPheLeuLeuSerIleThrThrGlyValHisSerGln 20
Db 13 ATGGAATGGAGTTGGATATTCTTCTCTCTCAGGAACCTGCAGGTGCCACTCTGAG 72

Qy 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 73 GTCCAGCTGCAGCAGCTCTGGACCTGAGCTGATAAGCCTGGGGCTTCAGTGAAGATGCC 132

Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMetHisTrpValLysGlnThrPro 60
Db 133 TGCAAGGCTCTGGATACACATTCTATGCTATGCTATGCTGGTGAACGACGAGCCT 192

Qy 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrAsn 80
Db 193 GGGCAGGCGCTTGATGGATGATATTAATCTTACATGATGCTACTAAGTACAAT 252

Qy 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerThrAlaTyrMet 100
Db 253 GAGAGGTCAAGGACGACCTGACTTACACAAATCTCCAGCAGCCTACATG 312

Qy 101 GlnIleSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120
Db 313 GAGCTCAGCAGCCTGACCTCTGAGGAGCTCTCGGCTCTATTACTGTGCAAGAGGACTTAT 372

Qy 121 -----GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSer 137
Db 373 TACTACGGTAGTAGGATTGACTACTGGGGCCAGGACCACTCTCAGTCTCTCTCA 432

Qy 138 AlaLysThrThrProProValTyrProLeuValProGlySer 152
Db 433 GCCAAAGCACACCCCATCTGCTATCCACTGGCCCCCTGGATCT 477

RESULT 11
MUSIGBIH1
LOCUS Mus musculus mRNA for immunoglobulin gamma-3 heavy chain precursor,
DEFINITION complete cds.
ACCESSION D14625
VERSION D14625.1 GI:286089
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1544)
AUTHORS Takahashi,S., Itoh,J., Nose,M., Ono,M., Yamamoto,T. and Kyogoku,M.
TITLE Cloning and cDNA sequence analysis of nephritogenic monoclonal
antibodies derived from an MRL/lpr lupus mouse
JOURNAL Mol. Immunol. 30 (2), 177-182 (1993)
MEDLINE 93156722
PUBMED 8429833
REFERENCE 2
AUTHORS Ono,M., Yamamoto,T., Kyogoku,M. and Nose,M.
TITLE Sequence analysis of the germ-line VH gene corresponding to a
nephritogenic antibody in MRL/lpr lupus mice
JOURNAL Clin. Exp. Immunol. 100 (2), 284-290 (1995)

MEDLINE 95262286
PUBMED 7743668
REFERENCE 3 (bases 1 to 1544)
AUTHORS Nose,M.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-1993) Masato Nose, Tohoku University School of
Medicine, Dept. of Pathology; 2-1 Seiry-cho Aoba-ku, Sendai,
Miyagi 980, Japan (E-mail:d22181@cctu.cc.tohoku.ac.jp,
Tel:81-22-273-9042, Fax:81-22-234-1986)
FEATURES
Location/Qualifiers
source 1..1544
/organism="Mus musculus"
/mol_type="mRNA"
/strain="MRL/MpJ-lpr/lpr"
/db_xref="taxon:10090"
/chromosome="14"
/clone="pB1H1"
/clone_lib="pbluescript KS(+)"
/dev_stage="adult"
1..32
33..1445
/note="precursor of C and V-D-J regions from 7B6.8"
/codon_start=1
/product="immunoglobulin gamma-3 heavy chain precursor"
/protein_id="BAA03476.1"
/db_xref="GI:1304160"
/translation="MEWSGVIFLLSVTAGVHQSQVQLQSGAELVRRPETSVMKSCKAA
GYTNTWIGVKORPGHLEWIGDIYPGGYTNVNEKFKATLTADTSSTAYMOL
SLLTSEDALYICARLGGYVYAMDYGGTSTVSSATTPASVYPLVPGCSDTSG
SSVTLGCLVKGIFPEFVTVKWNYSALSGVTVSSVLSLTVPSSTWPSQ
TVICNVAPASKTELIRKIEPRIPKSTPPGSSCPPNIGGSPVFIIPPAPKDALMI
SLTPKTCVVVDSEDDPDVHVSFVNDKVEHTAQTQPREAQYNSTFRVSALPQHQ
DMRGKEFKCKNNKALPAPIERTISKPKRAQTPQVTTIPPREQMSKKVSLTCLV
TNPTSEALISVSEWERNGELEODYKNTPTILSDGTIFYLYSKLTVDTSDWLGGEIFTSV
VHEAHNHHTQKLSRSPGK"
33..89
sig_peptide
mat_peptide
V_region 90..1442
/product="immunoglobulin gamma-3 heavy chain"
90..452
/product="immunoglobulin V-D-J region"
/note="7B6.8"
C_region 453..1442
/product="immunoglobulin gamma-3 secreted C-region"
3'UTR 1446..1544
polyA_signal 1514..1519
BASE COUNT 402 a 448 c 368 g 326 t
ORIGIN
Alignment Scores:
Pred. No.: 4.08e-52 Length: 1544
Score: 599.50 Matches: 116
Percent Similarity: 84.42% Conservative: 14
Best Local Similarity: 75.32% Mismatches: 21
Query Match: 73.47% Indels: 3
DB: 10 Gaps: 1

US-08-836-455-4 (1-153) x MUSIGBIH1 (1-1544)

Qy 1 MetGluCysSerTrpValPheLeuLeuSerIleThrThrGlyValHisSerGln 20
Db 33 ATGGAATGGAGCGGGTCTTTATCTTCTCTCTCAGTAACCTGCAGGTGCCACTCCAG 92

Qy 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 93 GTCCAACTGCAGCAGCTCTGGAGCTGAGCTGGTAAAGCCCTGGACTTCAGTGAAGATGCC 152

Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMetHisTrpValLysGlnThrPro 60
Db 153 TGCAGGCTCTGGATACACCTTCTACTACTGATAGGTGGGTAAAGCAGAGGCT 212

Qy 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrAsn 80
Db 213 GGACATGGCCCTTGAGTGGATGGAGATATTTACCCCTGGAGGTGGTTATATACTAACAAT 272

```





```

Pred. No.: 3,22e-52 Length: 489
Score: 595.00 Matches: 117
Percent Similarity: 81.65% Conservative: 12
Best Local Similarity: 74.05% Mismatches: 19
Query Match: 72.92% Indels: 10
DB: 10 Gaps: 2

US-08-836-455-4 (1-153) x MMU39900 (1-489)

Qy 1 MetGlucySerTrpValPheLeuLeuSerIleThrGlyValHisSerGln 20
Db 1 ATGAATGACGCTGGGTCTTCTCTCTCTCAATAACTCAGGTGCCATTGCCAG 60
Qy 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 61 GTCCAGCTGCAGCAGCTGACGCTGGAGCTGGTGAAGCCTGGGGCCTCAGTGAAGATTCC 120
Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstTrpValLysGlnThrPro 60
Db 121 TGCACAGCTTCTGGCTACACATTGAGTGGCTCTTGGATGAACCTGGGTGAAGCAGAGCCCT 180
Qy 61 GlyGlnGlyLeuGlnTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrAsn 80
Db 181 GGCAGGGCTTTCAGTGGATGGACGGATTATCTCTGGAGATGAGATTAATACTACAGT 240
Qy 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100
Db 241 GGAATAATTCAGGCAAGGCGACACTGACCTGCAGACAAATCCCTCCAGCAGAGCCTACTTG 300
Qy 101 GlnIleSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120
Db 301 CAATCAGCAGCTGACCTCTGTGGACTTCTGGCTCTATTTCTGTGCAAGAGAT----- 354
Qy 121 GluGly-----AlaLeuAspTyrTrpGlyGlnGlyThrSer 132
Db 355 GATGGTTACTACGTGCTTTGATTACTATGCTATGACATATGGGTCAAGGAACCTCA 414
Qy 133 ValThrValSerSerAlaLysThrThrProProValTyrProLeuValPro 150
Db 415 GTCACCGTCTCTCAGCAAAACAAACACCCCATCAGTCTATCCCTGCTGTTCTCT 468

RESULT 14
AB050070 469 bp mRNA linear ROD 02-APR-2002
LOCUS Mus musculus VH6G9 mRNA for anti-dsRNA (RDV-RNA) antibody, partial cds.
DEFINITION AB050070.1 GI:19909916
ACCESSION AB050070.1
VERSION AB050070.1
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Kitagawa,Y., Matsumoto,T., Okuhara,E. and Shikata,E.
TITLE Immunogenicity of rice dwarf virus-ribonucleic acid
JOURNAL Tohoku J. Exp. Med. 122 (4), 337-343 (1977)
MEDLINE 78034786
PUBMED 918970
REFERENCE 2
AUTHORS Kitagawa,Y.
TITLE anti-dsRNA (RDV-RNA) Ab VH region-VH6G9
JOURNAL Published Only in Database (2002)
REFERENCE 3 (bases 1 to 469)
AUTHORS Kitagawa,Y.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2000) Yoshichika Kitagawa, Akita Prefectural University, Biotechnology Institute; minami 2-2, Ogata, Akita 010-0444, Japan (E-mail:kitagawa@agri.akita-pu.ac.jp, URL:www.akita-pu.ac.jp/, Tel:81-185-45-2026(ex.400), Fax:81-185-45-2678)
FEATURES
source Location/Qualifiers
1. .469

```

/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 /cell\_line="hybridoma 6G9"  
 /cell\_type="B cell"  
 gene 1. .469  
 CDS /gene="VH6G9"  
 9. .>467  
 /gene="VH6G9"  
 /codon\_start=1  
 /product="anti-dsRNA (RDV-RNA) antibody"  
 /protein\_id="BAB87183.1"  
 /db\_xref="GI:19909917"  
 /translation="MKCSWILLFCQVTSQVQLQSGAEIARPGASVKLSKRA  
 GTFSSHWHGKQRPQGGIYPGDSDRYNKKFKGASLTADKSSSTAYIQLS  
 SLASEDSAVYICARNGNFDPDPDGQGLTVTSAAKTTPPPVPLAPGSL"

BASE COUNT 111 a 115 c 123 g 120 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1e-51 Length: 469  
 Score: 590.00 Matches: 113  
 Percent Similarity: 83.33% Conservative: 17  
 Best Local Similarity: 72.30% Mismatches: 20  
 Query Match: 72.30% Indels: 6  
 DB: 10 Gaps: 3

US-08-836-455-4 (1-153) x AB050070 (1-469)

Qy 1 MetGlucySerTrpValPheLeuLeuSerIleThrGlyValHisSerGln 20  
 Db 9 ATGAATGACGCTGGGTCTTCTCTCTCTCAATAACTCAGGTGCCATTGCCAG 68  
 Qy 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
 Db 69 GTTCAGCTCCAGCAGCTCTGGGCTGAGTGGCAAGACCTGGGCTTCAGTGAAGTTGTC 128  
 Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstTrpVal-----LysGln 58  
 Db 129 TGAAGGCTTCTGGCTACACCTTTAGTAGT-----CACTGGATACATGTTAAACAG 179  
 Qy 59 ThrProGlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyr 78  
 Db 180 AGCCTGGACAGGGCTCGAGTGGATGGGCCATTTATCTGGAGATGGTCTTAGG 239  
 Qy 79 TyrAsnGlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerThrAla 98  
 Db 240 TACAATAAGAAGTTCAAGGGCAAGGCTCATTCAGTGCAGATAAATCTCCAGCAGACC 299  
 Qy 99 TyrMetGlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArg--- 117  
 Db 300 TACATACAACTCAGCAGCTTGGCATCTGAGGACTCTCGGTCTATTACTGTGCAGGAAT 359  
 Qy 118 GlyAsnTrpGluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSer 137  
 Db 360 GGTAACTTCGACCCCTTTTGTATTATTATTTGGGCCAAGGACTCTGTGCTCTCTGCA 419  
 Qy 138 AlaLysThrThrProProValTyrProLeuValProGlySerLeu 153  
 Db 420 GCCAAAACAGCAGACCCCGCTCTATCTATCTATCTGCGCTTGGAGCTTG 467

RESULT 15  
 BC003435 1557 bp mRNA linear ROD 16-APR-2003  
 LOCUS Mus musculus immunoglobulin heavy chain 4 (serum IgG1), mRNA (cdna  
 DEFINITION clone MGC:6486 IMAGE:2646928), complete cds.  
 ACCESSION BC003435  
 VERSION BC003435.1 GI:13097380  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



REFERENCE  
AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1557)  
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butterfield, Y.S., Krzywinski, M.I., Jones, S.J., Skalska, U., Smalios, D.E.,  
 Schnerker, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22389257  
 12477932

2 (bases 1 to 1557)  
 Strausberg, R.  
 Direct Submission  
 Submitted (20-FEB-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-femail.nih.gov](mailto:cgabs-femail.nih.gov)  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,  
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 5 Row: 1 Column: 17.

FEATURES  
source

Location/Qualifiers  
 1. .1557  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="MGC:6486 IMAGE:2646928"  
 /tissue\_type="Mammary tumor. Metallothionien-TGF alpha  
 model. 10 month old virgin mouse. Taken by biopsy."  
 /clone\_lib="NCI-CGAP\_Mam1"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 1. .1557  
 /gene="Igh-4"  
 /note="synonym: IgG1"  
 /db\_xref="LocusID:16017"  
 /db\_xref="MGI:96446"  
 50. .1441  
 /codon\_start=1  
 /product="Igh-4 protein"  
 /protein\_id="AAH03435.1"  
 /db\_xref="GI:13097381"  
 /db\_xref="LocusID:16017"

## gene

## CDS

/translation="MEWIWIFLSTAGVHSQVQLQSQAGAEARPGASVRLSCKAS  
 GYFTGYGVSWVKQRTGQGLEWVYSGNTYYSEKFKGKATLTTDKSSYAYMHL  
 SLLTSEDVAIVCFQARSYYSDLFAYWGQGLTVVSAKTPPSVYPLAPGSAQTNS  
 VTNLGLCVKGYFEPETVTVNWSGLSSGVHTFPAVLQSDLYTLSSSVYVPSSTWPSST  
 MVRVHPASSIKVDFKIVPRDCGCRPCTCTPEVSSVFIFPKPKDLVILITLTPKVT  
 CVVYDSDKDEPVOFSWFDVDEVHTAQTRREEQNSTFRSEVSLPIMQDMLNGKE  
 FKRVNSAAFPAPETIKTISKGRKPAQVYITPPPKQMAKDKVSLTCTMTDFFPED  
 ITVWNGQPAENYKNTOPIMDTGDSYFIYSLNVQKSNWEAGNTFTCSVLHGLEHN  
 HHTKSLSHSPGK"  
 BASE COUNT 390 a 456 c 374 g 337 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.29e-50 Length: 1557  
 Score: 581.00 Matches: 109  
 Percent Similarity: 81.17% Conservative: 16  
 Best Local Similarity: 70.78% Mismatches: 27  
 Query Match: 71.20% Indels: 2  
 DB: 10 Gaps: 1  
 US-08-836-455-4 (1-153) x BC003435 (1-1557)  
 Qy 1 MetGluCysSerTrpValPheLeuPheLeuSerIleThrThrGlyValHisSerGln 20  
 Db 50 ATGGAATGGATCTGGATCTTTCTTCATCCTGTCAGGAACATGCAGGTGTCACCTCCAG 109  
 Qy 21 AlaTyrLeuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
 Db 110 GTTCAGCTGCAGCAGCTGTGGAGTGCAGGAGCCAGGGGCTTCAGTGAGGCTGTGCC 169  
 Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMetHisTrpValLysGlnThrPro 60  
 Db 170 TGCAAGGCTTCGGATACACCTTCACAGGCTATGGTCTAAGCTGGTGAAGCAGAGAAT 229  
 Qy 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80  
 Db 230 GGACAGGGCCTTGAGTGGTGGAGAGATTATCTCGAAGTGGCAATCTACTACAGT 289  
 Qy 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100  
 Db 290 GAGAAATTCAAGGCAAGGCCACACTACTACAGACAAATCCTCCACAGCAGCCTACATG 349  
 Qy 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120  
 Db 350 CACCTCAGCAGCCTGCATCTGAGGACTCTGAGGACTCTGAGTCTVATTCTGTGCAAGAAGCAGTTAC 409  
 Qy 121 GluGly-----AlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerAla 138  
 Db 410 TATAGTTACGACCTGTTTGTCTTACTGGGGCCAGGGGACTCTGGTCACTGTCTCTGTGCC 469  
 Qy 139 LysThrThrProProValTyrProLeuValProGlySer 152  
 Db 470 AAACGACACCCCATCTGTCTATCCACTGGCCCTGGATCT 511

Search completed: August 31, 2003, 00:02:04  
 Job time : 3133.85 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 30, 2003, 22:10:01 ; Search time 230.013 Seconds  
(without alignments)  
1795.608 Million cell updates/sec

Title: US-08-836-455-4

Perfect score: 816

Sequence: 1 MEC5WFLFLSITGVHSQ.....TVSSAKTTPPPVPLVPGSL 153

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US08836455/runat\_29082003\_132901\_22271/app\_query.fasta\_1.654  
-DB=N-Geneseq\_19Jun03 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THRESHOLD=100 -THRESHOLD\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US08836455 -CGN\_1\_1\_401\_@runat\_29082003\_132901\_22271 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N-Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	816	100.0	461	18	AAT85150 Murine monoclonal
2	816	100.0	461	20	AAT85150 hea
3	816	100.0	461	25	AAU51274 Mouse 11D10 antibo
4	621	76.1	1553	16	AAQ79930 Anti-tobacco mosai
5	602	73.8	540	10	AAQ91645 Heavy chain of mon
6	599	73.4	736	25	ABX16571 Mouse DNA encoding
7	574.5	70.4	1570	12	AAQ12637 Monoclonal antiod
8	570	69.9	725	21	AAQ43821 Mouse secreted exp
9	568	69.6	1683	20	AAQ61084 Mouse immunoglobul
10	565	69.2	458	10	AAQ91146 2H7 Vh sequence.
11	565	69.2	459	20	AAQ82337 Mouse antibody 2H7
12	565	69.2	459	22	AAH22070 2H7 heavy chain va
13	565	69.2	491	18	AAQ70868 Mouse 2H7 antibody
14	565	69.2	491	19	AAQ18557 Mouse 2H7 antibody
15	565	69.2	491	19	AAQ03926 Mouse 2H7 antibody
16	565	69.2	491	19	AAQ18593 Mouse 2H7 antibody
17	565	69.2	518	18	AAT36316 2H7 antibody heavy
18	565	69.2	520	18	AAT51042 Coding sequence fo
19	563	69.0	1392	18	AAT62936 3F4 (Chimeric) hum
20	563	69.0	1395	18	AAT62935 Murine anti-porcine
21	563	69.0	1581	14	AAQ48037 Monoclonal antiod
22	563	69.0	3400	18	AAT62937 3F4 human C2/G4 ch
23	563	69.0	5300	18	AAT62938 3F4 human IgG4 exp
24	561.5	68.8	1083	24	AAQ45753 Human CH2 domain d
25	561.5	68.8	1083	25	ABQ24017 Antibody C2B8 CH2
26	561.5	68.8	1413	21	AAQ63531 DNA encoding a dim
27	561.5	68.8	1413	24	AAQ45752 Human C2B8 antiod
28	561.5	68.8	1413	25	ABQ24016 Antibody C2B8 heav
29	561.5	68.8	9208	15	ABQ65629 Vector coney. TCAE
30	561.5	68.8	9209	25	ABQ94203 TCAE8 expression v
31	561.5	68.8	19001	19	AAQ61793 Traget plasmid Mol
32	560.5	68.7	1392	19	AAQ70129 Anti-Fas MAB HFE7A
33	560.5	68.7	1392	21	AAQ72108 cDNA encoding mous
34	560.5	68.7	1392	21	AAQ11546 Murine anti-Fas an
35	560.5	68.7	1392	24	ABL45925 Mouse humanised an
36	560.5	68.7	1392	24	ABL48668 Humanised anti-Fas
37	555	68.0	420	22	AAF81910 Anti-CA125 bifunct
38	554.5	68.0	1314	24	AAQ29310 Human penton base
39	554.5	68.0	1314	25	ABX12746 DNA encoding porti
40	554.5	68.0	1516	24	AAQ29308 Human penton base
41	554.5	68.0	1516	25	ABX12744 DNA encoding mouse
42	554.5	68.0	1582	16	AAQ94037 MAB 55.1 heavy cha
43	552.5	67.7	409	21	AAQ87715 Anti-human VEGF re
44	552.5	67.7	409	22	AAQ70192 Anti-human Flt-1 m
45	552	67.6	468	19	AAQ05544 Nucleotide sequenc

ALIGNMENTS

RESULT 1	
AAT85150	
ID	AAT85150 standard; CDNA; 461 BP.
XX	
AC	AAT85150;
XX	
DT	25-MAR-2003 (updated)
DT	04-JAN-1998 (first entry)
XX	
DE	Murine monoclonal anti-idiotyp antibody 11D10 VH cDNA.
XX	
KW	Monoclonal antibody 11D10; anti-idiotyp antibody; mucin;
KW	human milk fat globule; HMF; tumour; breast cancer; vaccine; ss.
XX	
OS	Mus musculus.
XX	
FH	Key Location/Qualifiers



Pred. No.: 1.53e-70 Length: 461  
Score: 816.00 Matches: 153  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0

US-08-836-455-4 (1-153) x AAV83773 (1-461)

QY 1 MetGluCysSerTrpValPheLeuPheLeuSerIleThrThrGlyValHisSerGln 20  
|||||  
Db 1 ATGGAATGACGTGGGTCTTCTCTCTCTCAATAACTACAGGTGCCACTCCAG 60  
QY 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
|||||  
Db 61 GCTTATCTACAGCAGCTGGGGCTGAGTGTGAGTCTGGGGCTCAGTGAAGATGCC 120  
QY 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstTrpValLysGlnThrPro 60  
|||||  
Db 121 TGCAGAGCTTCTGGCTACACATTGACCAAGTTACAAATATGCACCTGGGTAAAGCAGACACCT 180  
QY 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrAsn 80  
|||||  
Db 181 GGACAGGCTCGAATGGATGGAAATATTTTCTGGAATGGTGATCTTACTACAAAT 240  
QY 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100  
|||||  
Db 241 CAGAAGTTTAAAGGCAAGGCTCATTTGACTGCAGACACATCTCCAGCAGCCTACATG 300  
QY 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120  
|||||  
Db 301 CAGATCAGCAGCTGACATCTGAAGACTCTGCGGTCTATTCTGTGCAAGAGGGAACCTGG 360  
QY 121 GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThr 140  
|||||  
Db 361 GAGGGTCTCTGACTACTGGGGTCAAGGAACCTCAGTACCCTGCTCTCTCCTCAGCCAAACG 420

RESULT 3

ID AAL51274 standard; cDNA; 461 BP.

XX AC AAL51274;

XX 20-MAR-2003 (first entry)

XX Mouse 11D10 antibody heavy chain variable region coding sequence.

XX Mouse; murine; gene; ss; vaccine; tumour; human milk fat globules; HMFG;

XX carcinoembryonic antigen; CEA; 11D10; 3H1; HMFG-associated tumour;

XX CEA-associated tumour; anti-idiotype antibody.

XX Mus musculus.

XX Key Location/Qualifiers

XX CDS 1..459

XX /tag= a

XX /partial

XX /product= "Mouse 11D10 anti-idiotype antibody heavy chain

XX variable region"

XX /note= "No stop codon is given"

XX sig\_peptide 1..57

XX /tag= b

XX mat\_peptide 58..459

XX /tag= c

XX WO200292012-A2.

XX 21-NOV-2002.

XX

PF 17-MAY-2002; 2002WO-US15840.

XX PR 17-MAY-2001; 2001US-0861294.

XX PA (KENT ) UNIV KENTUCKY RES FOUND.

XX PI Chatterjee M, Foon KA;

XX WIPI; 2003-129216/12.

XX DR P-PSDB; AAO16293.

XX Use of anti-idiotypic antibodies for human milk fat globules (HMFG)- or

XX PT carcinoembryonic antigen (CEA)-associated tumor for delaying the

XX PT development of, or treating a HMFG- or CEA-associated tumor (e.g.

XX PT breast tumor) in humans

XX PS Disclosure; Fig 2; 98pp; English.

XX CC The invention comprises a method for delaying the development of, or

XX CC treating a tumor that is associated with human milk fat globules (HMFG)

XX CC or carcinoembryonic antigen (CEA). The method of the invention involves

XX CC administering an anti-idiotypic antibody for HMFG (e.g. 11D10), and an

XX CC anti-idiotypic antibody for CEA (e.g. 3H1). The method is useful for

XX CC delaying the development, of or treating HMFG/CEA-associated tumours. The

XX CC present DNA sequence encodes the heavy chain variable region of the mouse

XX CC 11D10 anti-idiotype antibody.

XX SQ Sequence 461 BP; 112 A; 121 C; 119 G; 109 T; 0 other;

Alignment Scores:

Pred. No.: 1.53e-70 Length: 461

Score: 816.00 Matches: 153

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 25 Gaps: 0

US-08-836-455-4 (1-153) x AAL51274 (1-461)

QY 1 MetGluCysSerTrpValPheLeuPheLeuSerIleThrThrGlyValHisSerGln 20

Db 1 ATGGAATGACGTGGGTCTTCTCTCTCTCAATAACTACAGGTGCCACTCCAG 60

QY 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40

Db 61 GCTTATCTACAGCAGCTGGGGCTGAGTGTGAGTCTGGGGCTCAGTGAAGATGCC 120

QY 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstTrpValLysGlnThrPro 60

Db 121 TGCAGAGCTTCTGGCTACACATTGACCAAGTTACAAATATGCACCTGGGTAAAGCAGACACCT 180

QY 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTrpAsn 80

Db 181 GGACAGGCTCGAATGGATGGAAATATTTTCTGGAATGGTGATCTTACTACAAAT 240

QY 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100

Db 241 CAGAAGTTTAAAGGCAAGGCTCATTTGACTGCAGACACATCTCCAGCAGCCTACATG 300

QY 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120

Db 301 CAGATCAGCAGCTGACATCTGAAGACTCTGCGGTCTATTCTGTGCAAGAGGGAACCTGG 360

QY 121 GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThr 140

Db 361 GAGGGTCTCTGACTACTGGGGTCAAGGAACCTCAGTACCCTGCTCTCTCCTCAGCCAAACG 420

QY 141 ThrProProValTyrProLeuValProGlySerLeu 153

Db 421 ACACCCCCACCCGCTTATCCACTGGTCCCTGGAAGCTTG 459

RESULT 4

AAQ79930

AAQ79930 standard; cDNA to mRNA; 1553 BP.  
 AAQ79930;  
 01-SEP-1995 (first entry)  
 Anti-tobacco mosaic virus monoclonal Ab heavy chain cDNA.  
 Tobacco mosaic virus; TMV; monoclonal antibody;  
 heavy chain; virus-resistant plants; biofarming; ss.  
 Synthetic.  
 Key Location/Qualifiers  
 CDS 56..1453  
 FT /\*tag= a  
 FT sig\_peptide 56..112  
 FT /\*tag= b  
 FT mat\_peptide 113..1450  
 FT /\*tag= c  
 JP06319396-A.  
 22-NOV-1994.  
 07-MAY-1993; 93JP-0131208.  
 07-MAY-1993; 93JP-0131208.  
 (NLSB ) JAPAN TOBACCO INC.  
 (KURS ) KURARAY CO LTD.  
 WPI; 1995-040220/06.  
 P-PSDB; AAR66758.  
 Transformed plant producing animal-derived anti-virus antibody -  
 esp. tobacco plants producing anti-tobacco mosaic virus  
 monoclonal antibody  
 Example 2; Pages 14-15; 26pp; Japanese.  
 AAQ79929 and AAQ79930 encode AAR66757 and AAR66758, the light and heavy  
 chains of an animal derived anti-tobacco mosaic virus (TMV)  
 monoclonal antibody. The cDNAs were incorporated into a T1  
 plasmid vector, which was incorporated into A. tumefaciens.  
 The resultant plant expression vector was used to transform  
 tobacco plants, making them TMV resistant, the plants could  
 also be biofarmed for the prodn. of anti-virus antibodies.  
 Sequence 1553 BP; 386 A; 450 C; 372 G; 345 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 5,92e-51 Length: 1553  
 Score: 621.00 Matches: 118  
 Percent Similarity: 85.26% Conservative: 15  
 Best local Similarity: 75.64% Mismatches: 19  
 Query Match: 76.10% Indels: 4  
 DB: 16 Gaps: 2  
 US-08-836-455-4 (1-153) x AAQ79930 (1-1553)  
 Qy 1 MetGluCysSerTrpValPheLeuPheLeuSerIleThrThrGlyValHisSerGln 20  
 Db 56 ATGGAATGTAACCTGATACCTCCCTTTATCTGTGTCAGTAACCTCAGGTGCTACTACAG 115  
 Qy 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
 Db 116 GTTCAGCTCCAGCAGTCTGGGGCTGAGCTGGCAAGACCTGGGGCTTCAGTGAAGTTGCC 175  
 Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstTrpValLysGlnThrPro 60  
 Db 176 TGCAAGGCTTCTGGCTACACCTTTTACTAGCTAGCTAGTGGGTAAACACAGAGGCGCT 235  
 Qy 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80  
 Db 236 GGACAGGGTCTGGAATGGATTGGGGCTATTATCTCTGGAATGGTATAGTACACT 295  
 Qy 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100  
 Db 296 CAGAAGTTCAAGGGCAAGGCCACATTCAGCTGCAGATAAATCTCCACACAGACCTACATG 355  
 Qy 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGly----- 118  
 Db 356 CAACTCAGCGCTTGGCATCTGAGGACTCTGGGCTCTATTACTGTGCAAGAGAGGGGGT 415  
 Qy 119 ---AsnTrp---GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSer 136  
 Db 416 TACTCTGGTCCGACTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCC 475  
 Qy 137 SerAlaLysThrThrProProValTyrProLeuValProGlySer 152  
 Db 476 TCAGCCAAAACAGACACCCCATCTGTCTATCCACTGGCCCTGGATCT 523  
 RESULT 5  
 AAN91645  
 ID AAN91645 standard; DNA; 540 BP.  
 XX  
 AC AAN91645;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 31-OCT-2002 (updated)  
 DT 15-MAR-1990 (first entry)  
 XX  
 DE Heavy chain of monoclonal antibody 6A4.  
 XX  
 KW Monoclonal antibody 6A4; heavy chain; Pseudomonas aeruginosa; OMP-1.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 CDS 64..540  
 FT /\*tag= a  
 FT  
 FT  
 FT  
 PN EP338395-A.  
 XX  
 PD 25-OCT-1989.  
 XX  
 PF 12-APR-1989; 89EP-0106463.  
 XX  
 PR 19-APR-1988; 88DE-3813023.  
 XX  
 PA (BEHW ) BEHRINGWERKE AG.  
 XX  
 PI Domdey H, Marget M, Vonspecht B;  
 XX  
 DR WPI; 1989-310861/43.  
 DR P-PSDB; AAP93079.  
 XX  
 PT Monoclonal antibody to Pseudomonas aeruginosa - and DNA coding for  
 variable antibody regions.  
 XX  
 PS Claim 1; page 6; 7pp; german.  
 XX  
 CC The sequence has a variable and constant region. Monoclonal antibody 6A4  
 CC reacts with the OMP-1 protein of all 19 known serotypes of P.aeruginosa.  
 CC It is used for therapy and diagnosis of infection, and as a carrier for  
 CC drugs. The antibody is IgG2a subclass.  
 CC (Updated on 31-OCT-2002 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 540 BP; 145 A; 144 C; 130 G; 121 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1.17e-49 Length: 540  
 Score: 602.00 Matches: 116



```

RESULT 7
AAQ12637
ID AAQ12637 standard; DNA; 1570 BP.
XX
AC
AC
XX
XX
DT 25-MAR-2003. (updated)
DT 03-OCT-1991 (first entry)
XX
XX Monoclonal antibody OK3T heavy chain coding sequence.
XX
XX OK3T; heavy chain; humanised antibodies; CDR-grafting; ss.
XX
OS Mus musculus.
XX
XX Key Location/Qualifiers
FH sig_peptide 41..97
FT /*tag= a
FT mat_peptide 98..1447
FT /*tag= b
FT /product= OK3T heavy chain
XX
XX W09109967-A.
PN W09109968-A.
XX
XX 11-JUL-1991.
XX
XX 21-DEC-1990; 90WO-GB02017.
XX
XX 21-DEC-1990; 90WO-GB02017.
PR 21-DEC-1989; 89GB-0028874.
XX
XX (CLUT ) CELLTech LTD.
XX
XX Adair JR, Athwal DS, Emtage JS;
XX
XX WPI; 1991-222915/30.
DR P-PSDB; AARI3061.
XX
XX New humanised antibodies comprising CDR grafted antibody - with
XX heavy and light chains, for use in in vivo therapy and diagnosis
XX
XX Disclosure; Fig 2a; 91pp; English.
XX
XX The OK3T heavy chain sequence was isolated from a cDNA library
XX prepared from OK3T producing cells. The library was screened with a
XX probe complementary to a sequence in the mouse IgG2a constant CH1
XX domain region. The OK3T sequence was used in CDR-grafting experiments
XX to prepare humanised antibodies.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX SQ Sequence 1570 BP; 447 A; 443 C; 356 G; 324 T; 0 other;

Alignment Scores:
Pred. No.: 1.99e-46 Length: 1570
Score: 574.50 Matches: 113
Percent Similarity: 80.13% Conservative: 8
Best Local Similarity: 74.83% Mismatches: 29
Query Match: 70.40% Indels: 1
DB: 12 Gaps: 1

US-08-836-455-4 (1-153) x AAQ12637 (1-1570)
Qy 1 MetGluCysSerTrpValPheLeuPheLeuSerIleThrThrGlyValHisSerGln 20
Db 41 ATGGAAGGCACTGGATCTTCTACTCTCTGTGTGTCAGTAAGTGCAGGTGCCACTCCAG 100
Qy 21 AlaTyrLeuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 101 GTCCAGCTGCAGCAGTCTGGGGCTGAACCTGGCAAGACCTGGGGCTCAGTGAAGATGCC 160
Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMetHisTrpValLysGlnThrPro 60

```

```

Db 161 TGCAAGGCTTCTGGCTACACCTTTACTAGGTACACGATGCACCTGGGTAAACAGAGCCT 220
Qy 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80
Db 221 GGACAGGCTCTGGATGGATGGATACATTAATCTAGCGCTGGTTATACTAATACAAT 280
Qy 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100
Db 281 CAGAAGTTCAGGACAAAGGCCACATTCGACTACAGACAAATCCTCCAGCAGCCTACATG 340
Qy 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArg--GlyAsn 119
Db 341 CAACTGAGCAGCCTGACATCTGAGGACTCTGAGTCTATTACTGTGAAGATATATGAT 400
Qy 120 TrpGluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLys 139
Db 401 GATCATTACTGCCTTGACTACTGGGGCCAGGCCACCACTCTCAGCTCTCTCTCAGCCAAA 460
Qy 140 ThrThrProProValTyrProLeuValPro 150
Db 461 ACAACAGCCCATCGGTCTATCCACTGGCCCT 493

RESULT 8
AAA43821
ID AAA43821 standard; cDNA; 725 BP.
XX
AC
AC
XX
XX 21-AUG-2000 (first entry)
XX
XX Mouse secreted expressed sequence tag SEQ ID NO:396.
XX
XX Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;
XX expressed sequence tag; EST; probe; chemotactic; proliferative;
XX immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
XX thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
XX antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;
XX antitumor; osteopathic; neuroprotective; nootropic; antipsoriatic;
XX cerebroprotective; anticonvulsant; antidepressant; gene therapy;
XX vaccine; autoimmune diabetes; asthma; myeloid cell deficiency; ulcer;
XX lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
XX central nervous system disorder; Alzheimer's disease; stroke;
XX Parkinson's disease; Huntington's disease; coagulation disorder;
XX haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
XX tumour; infection; depression; psoriasis; ss.
XX
XX Mus musculus.
XX
XX WO200021991-A1.
XX
XX 20-APR-2000.
XX
XX 15-OCT-1999; 99WO-US24206.
XX
XX 15-OCT-1998; 98US-0104436.
XX
XX (GEMY ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
XX Merberg D, Treacy M, Bowman MR;
XX
XX WPI; 2000-317938/27.
XX
XX Isolated polynucleotides, and encoded proteins, comprising secreted
XX expressed sequence tags (SESTs), useful for treating various disorders
XX such as autoimmune, infectious, and central nervous system disorders -
XX
XX Claim 1; Page 306; 803pp; English.
XX
XX AAA43426 to AAA45925 represent specifically claimed secreted expressed
XX sequence tags (SESTs), isolated from human, mouse, chicken and rat
XX

```









[illegible]

DB: 19 Gaps: 1

US-08-836-455-4 (1-153) x AAV18557 (1-491)

1 MetGlucySsrTrpValPheLeuPheLeuSerIleThrGlyValHisSerGln 20  
 72 ATGGGATTACAGCAGGATCTTCTCTCCTCCTGCAGTACAGGTGTCCACTCCCG 131  
 21 AlaTyrLeuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
 132 GCTTATCTACAGCAGTCTGGGGCTGAGCTGGTGAGGCCCTCAGTGAAGATGTCC 191  
 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstPValLysGlnThrPro 60  
 192 TGCAGGGCTTCTGGCTACACATTTACCAGATTACAATATGCACCTGGGTAAAGCAGACACCT 251  
 61 GlyGlnGlyLeuGluTTPiLeGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80  
 252 ACACAGGGCCCTGGAAATGGATTGGAGCTATTATTCCAGGAATGGTGATCTTGGCTACAAT 311  
 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100  
 312 CAGAAGTTCAGGCCAAGGCCACACTCAGCTAGACAAATCCTCCAGCACAGCCTACATG 371  
 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArg----- 117  
 372 CAGCTCAGCAGCCTGACATCTGAAGACTCTCGGCTCTATTCTGTGCAAGAGTGGTGATC 431  
 118 --GlyAsnTrpGluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSer 136  
 432 TATAGTAACTCTTACTGGTACTTCGATGCTCGGGGCACAGGACCACGGTCACCGTCTCG 491

RESULT 15

AAV03926

ID AAV03926 standard; cDNA: 491 BP.

AC AAV03926;

01-JUN-1998 (first entry)

Mouse 2H7 antibody heavy chain variable region cDNA.

Mouse; murine; heavy chain; variable region;  
 immunoglobulin fragment production; Ig fragment production;  
 monoclonal antibody 2H7; human B-cell surface antigen; ss.

Mus sp.

Key	Location/Qualifiers
sig_peptide	72..128
mat_peptide	/*tag= a
FT	129..491
FT	/*tag= b
FT	
US5693493-A.	
02-DEC-1997.	
25-MAY-1995;	95US-0450731.
29-MAR-1990;	90US-0501092.
01-NOV-1985;	85US-0793980.
27-OCT-1986;	86WO-US02269.
24-JUL-1987;	87US-0077528.
11-JAN-1988;	88US-0142039.
08-DEC-1992;	92US-0987555.
18-AUG-1994;	94US-0239085.
25-MAY-1995;	95US-0450731.
(XOWA.) XOWA CORP.	
XX	
XX	
XX	
PI	Better M, Horwitz AH, Lei S, Liu AY, Robinson RR;
PI	Wall R, Wilcox GL;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 30, 2003, 22:11:56 ; Search time 202.802 seconds  
(without alignments)  
1730.853 Million cell updates/sec

Title: US-08-836-455-4

Perfect score: 816

Sequence: 1 MECSWFLFLSITGTHSQ.....TVSSAKTTPPVYPLVPGSL 153

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO.spool/US08836455/runat\_29082003\_132902\_22302/app\_query.fasta.1.654  
-DB=publishedApplications\_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US08836455 -CGN\_1\_1\_271 -runat\_29082003\_132902\_22302  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	816	100.0	461	9	US-09-861-294-3
					Sequence 3, Appli

2	816	100.0	461	12	US-10-367-506-3	Sequence 3, Appli
3	599	73.4	736	13	US-10-006-773-12	Sequence 12, Appl
4	565.5	69.3	1570	11	US-09-795-515-6	Sequence 6, Appli
5	561.5	68.8	9199	11	US-09-911-692-3	Sequence 3, Appli
6	561.5	68.8	9209	10	US-09-911-703-3	Sequence 2, Appli
7	561.5	68.8	9209	11	US-09-905-928-2	Sequence 2, Appli
8	561.5	68.8	9209	12	US-10-238-681-3	Sequence 2, Appli
9	561.5	68.8	9209	14	US-10-096-964-2	Sequence 2, Appli
10	561.5	68.8	18986	13	US-10-109-853-2	Sequence 2, Appli
11	560.5	68.7	1392	14	US-10-216-484-8	Sequence 8, Appli
12	554.5	68.0	1314	10	US-09-903-327A-5	Sequence 5, Appli
13	554.5	68.0	1516	10	US-09-903-327A-1	Sequence 1, Appli
14	552.5	67.7	409	14	US-10-160-232-7	Sequence 7, Appli
15	540	66.2	406	12	US-10-244-821-85	Sequence 85, Appl
16	540	66.2	406	14	US-10-150-762-85	Sequence 85, Appl
17	540	66.2	406	14	US-10-150-762-87	Sequence 87, Appl
18	539.5	66.1	540	9	US-09-748-960-5	Sequence 5, Appli
19	532.5	65.3	5691	10	US-09-897-006-11	Sequence 11, Appl
20	532.5	65.3	5691	11	US-09-897-511A-11	Sequence 11, Appl
21	531.5	65.1	482	9	US-09-881-823-19	Sequence 19, Appl
22	531	65.1	478	13	US-10-040-739-911	Sequence 911, App
23	530	65.0	418	11	US-09-269-921-2	Sequence 2, Appli
24	530	65.0	418	11	US-09-509-098-3	Sequence 3, Appli
25	530	65.0	418	15	US-10-218-253-2	Sequence 2, Appli
26	528	64.7	418	11	US-09-269-921-75	Sequence 75, Appl
27	528	64.7	418	11	US-09-509-098-97	Sequence 97, Appl
28	528	64.7	418	15	US-10-218-253-75	Sequence 75, Appl
29	522.5	64.0	420	11	US-09-905-928-5	Sequence 5, Appli
30	522.5	64.0	420	14	US-10-096-964-5	Sequence 5, Appli
31	521.5	63.9	1518	14	US-10-207-655-304	Sequence 304, App
32	521.5	63.9	1572	14	US-10-207-655-298	Sequence 298, App
33	521.5	63.9	1584	14	US-10-207-655-285	Sequence 285, App
34	520.5	63.8	420	10	US-09-911-703-9	Sequence 9, Appli
35	520.5	63.8	420	12	US-10-238-681-10	Sequence 10, Appl
36	520.5	63.8	453	11	US-09-911-692-9	Sequence 9, Appli
37	518	63.5	1524	12	US-10-053-530-5	Sequence 5, Appli
38	518	63.5	1524	14	US-10-207-655-5	Sequence 5, Appli
39	518	63.5	1536	14	US-10-207-655-283	Sequence 283, App
40	517	63.4	1521	14	US-10-207-655-351	Sequence 351, App
41	516.5	63.3	474	12	US-09-226-157-1	Sequence 1, Appli
42	516.5	63.3	4191	14	US-10-176-380-18	Sequence 18, Appl
43	514.5	63.1	1515	14	US-10-207-655-380	Sequence 380, App
44	514.5	63.1	1515	14	US-10-207-655-382	Sequence 382, App
45	514.5	63.1	1515	14	US-10-207-655-383	Sequence 383, App

#### ALIGNMENTS

RESULT 1  
US-09-861-294-3  
; Sequence 3, Application US/09861294  
; Patent No. US20020098190A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TUMORS BEARING HMG AND CEA ANTIGENS  
; FILE OF INVENTION: TUMORS BEARING HMG AND CEA ANTIGENS  
; FILE REFERENCE: 30414200620  
; CURRENT APPLICATION NUMBER: US/09/861,294  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 461  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(461)











GENERAL INFORMATION:  
APPLICANT: ANDERSON, DARRELL R.  
APPLICANT: HANNA, NABIL  
APPLICANT: LEONARD, JOHN E.  
APPLICANT: NEWMAN, ROLAND A.  
APPLICANT: REFF, MITCHELL E.  
APPLICANT: RASTETTER, WILLIAM H.  
TITLE OF INVENTION: THERAPEUTIC APPLICATION OF CHIMERIC AND RADIO LABELED ANTIBODIES TO HUMAN B LYMPHOCYTE RESTRICTED DIFFERENTIATION ANTIGEN FOR TREATMENT OF B CELL LYMPHOMA  
FILE REFERENCE: 37003/0291808  
CURRENT APPLICATION NUMBER: US/10/238,681  
CURRENT FILING DATE: 2002-09-11  
PRIOR APPLICATION NUMBER: 08/921,060  
PRIOR FILING DATE: 1997-08-29  
PRIOR APPLICATION NUMBER: 08/149,099  
PRIOR FILING DATE: 1993-11-03  
PRIOR APPLICATION NUMBER: 07/978,891  
PRIOR FILING DATE: 1992-11-13  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 3  
LENGTH: 9209  
TYPE: DNA  
ORGANISM: Murine sp.  
US-10-238-681-3

Alignment Scores:  
Pred. No.: 2,25e-59 Length: 9209  
Score: 561.50 Matches: 109  
Percent Similarity: 80.00% Conservative: 15  
Best Local Similarity: 70.32% Mismatches: 28  
Query Match: 68.81% Indels: 3  
DB: 12 Gaps: 1

US-08-836-455-4 (1-153) x US-10-238-681-3 (1-9209)

Qy 1 MetGluCysSerTrpValPheLeuSerIleThrGlyValHisSerGln 20  
Db 2401 ATGGTTGGAGCCCTCATCTTCTCTCTCTGCTGTGCTACGCGTCTCTGCCAG 2460

Qy 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
Db 2461 GTACAACCTGCAGCAGCCCTGGGGCTGAGCTGGTGAAGCCTGGGGCTCAGTGAAGATGTCC 2520

Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMetHisTrpValLysGlnThrPro 60  
Db 2521 TGCAGGCTTCTGGCTACACATTTTACAGTTACATATGCTGGTAAACACACACCT 2580

Qy 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80  
Db 2581 GGTCGGGCGCTGGAATGATGGAGCTATTTATCCGGAATGCTGATCTCTCTACAAT 2640

Qy 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerThrAlaTyrMet 100  
Db 2641 CAGAAGTTCAAGCAAGCCACATTTGCTGACACAAATCTCCAGCAGCAGCTACATG 2700

Qy 101 GlnIleSerLeuThrSerGluAspSerAlaValTyrPheCysAlaAatGlyAsnTrp 120  
Db 2701 CAGTCAGCAGCCTGACATCTGCGGTCTATTTACTGTGCAAGATCGACTTAC 2760

Qy 121 GluGlyAla-----LeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSer 137  
Db 2761 TACGGCGGTGACTGCTTCAATGCTCTGGGGCGCAGGACCGGTCAACGCTCTGTCA 2820

Qy 138 AlaLysThrThrProProValTyrProLeuValProGlySer 152  
Db 2821 GCTAGACCAAGGCGCCATCGGTCTCTCCCGCTGGCAGCCCTCTCC 2865

RESULT 9  
US-10-096-964-2  
; Sequence 2, Application US/10096964

Publication No. US20030082172A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Darrell R.  
Hanna, Nabil  
Leonard, John E.  
Newman, Roland A.  
Reff, Mitchell E.  
Rastetter, William H.  
TITLE OF INVENTION: Therapeutic Application of Chimeric and Radiolabeled Antibodies to Human B Lymphocyte Restricted Differentiation Antigen for the Treatment of B-Cell Lymphoma  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: 699 Prince St.  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/096,964  
FILING DATE: 14-Mar-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,813  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/149,099  
FILING DATE: 03-NOV-1993  
APPLICATION NUMBER: US 07/978,891  
FILING DATE: 13-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-158  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9209 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: anti-CD20 in TCAE 8  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-096-964-2

Alignment Scores:  
Pred. No.: 2,25e-59 Length: 9209  
Score: 561.50 Matches: 109  
Percent Similarity: 80.00% Conservative: 15  
Best Local Similarity: 70.32% Mismatches: 28  
Query Match: 68.81% Indels: 3  
DB: 14 Gaps: 1

US-08-836-455-4 (1-153) x US-10-096-964-2 (1-9209)

Qy 1 MetGluCysSerTrpValPheLeuSerIleThrGlyValHisSerGln 20  
Db 2401 ATGGTTGGAGCCCTCATCTTCTCTCTCTGCTGTGCTACGCGTCTCTGCCAG 2460

Qy 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
Db 2461 GTACAACCTGCAGCAGCCCTGGGGCTGAGCTGGTGAAGCCTGGGGCTCAGTGAAGATGTCC 2520

```
QY 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethHisTrpValLysGlnThrPro 60
Db 2521 TGAAGGCTTCTGGCTACACATTTACAGTTTACAATATGACCTGGGTAAACACACACCT 2580
QY 61 GlyGlnGlyLeuGluTrpLleGlyAsnLlePheProGlyAsnGlyAspThrTyrAsn 80
Db 2581 GTCGGGGCTGGAATGGATGAGCTATTATTCGCCGGAATGGTGATCTTCTTCAAT 2640
QY 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerThrAlaTyrMet 100
Db 2641 CACAAGTTCAAGCAAGGCACTTACCTGCTGCAACAATCTCCAGCACAGCTACATG 2700
QY 101 GlnLysSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120
Db 2701 CAGCTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATTACTGTGCAAGATCGACTTAC 2760
QY 121 GluGlyAla-----LeuAspTyrTrpGlyGlnGlyThrSerValThrValSer 137
Db 2761 TACGGGGCTGACTGGTACTTCAATGCTGGGGCGCAGGGACACGCTCACCCTCTGCA 2820
QY 138 AlaLysThrThrProProValTyrProLeuValProGlySer 152
Db 2821 GCTAGCACCAAGGCCCATCGGCTTCCCTCGGCACCTCCTCC 2865
RESULT 10
US-10-109-853-2
; Sequence 2, Application US/10109853
; Publication No. US20020192820A1
; GENERAL INFORMATION:
; APPLICANT: REFF, Mitchell E.
; BARNETT, Richard Spence
; MCLACHLAN, Karen Retta
; TITLE OF INVENTION: NOVEL METHOD FOR INTEGRATING GENES AT
; SPECIFIC SITES IN MAMMALIAN CELLS VIA HOMOLOGOUS
; RECOMBINATION AND VECTORS FOR ACCOMPLISHING THE SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/109,853
; FILING DATE: 01-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,715
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/819,866
; FILING DATE: 14-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 012712-352
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1896 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-109-853-2
Alignment Scores:
Pred. No.: 5,99e-59 Length: 1896
Score: 561.50 Matches: 109
Percent Similarity: 80.00% Conservative: 15
Best Local Similarity: 70.32% Mismatches: 28
Query Match: 68.81% Indels: 3
DB: 13 Gaps: 1

US-08-836-455-4 (1-153) x US-10-109-853-2 (1-18986)
QY 1 MetGluCysSerTrpValPheLeuPheLeuSerIleThrThrGlyValHisSerGln 20
Db 9433 ATGGGTTGGAGCCTCATCTTCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9492
QY 21 AlaTyrLeuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 9493 GTCAACTGCAGCAGCCTGGGCTGAGCTGGTGAAGCCTGGGGCTCAGTGAAGATGTCC 9552
QY 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethHisTrpValLysGlnThrPro 60
Db 9553 TGAAGGCTTCTGGCTACACATTTACCACTTACCAATATGCACTGGGTAAACACACACCT 9612
QY 61 GlyGlnGlyLeuGluTrpLleGlyAsnLlePheProGlyAsnGlyAspThrTyrTyrAsn 80
Db 9613 GGTCTGGGGCTGGAATGGATGAGCTATTATTCGCCGGAATGGTGATCTTCTTCAAT 9672
QY 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerThrAlaTyrMet 100
Db 9673 CAGAAGTTCAAGCAAGGCACTTACCTGCTGCAACAATCTCCAGCACAGCTACATG 9732
QY 101 GlnLysSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120
Db 9733 CAGCTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATTACTGTGCAAGATCGACTTAC 9792
QY 121 GluGlyAla-----LeuAspTyrTrpGlyGlnGlyThrSerValThrValSer 137
Db 9793 TACGGGGCTGACTGGTACTTCAATGCTGGGGCGCAGGGACACGCTCACCCTCTGCA 9852
QY 138 AlaLysThrThrProProValTyrProLeuValProGlySer 152
Db 9853 GCTAGCACCAAGGCCCATCGGCTTCCCTCGGCACCTCCTCC 9897

RESULT 11
US-10-216-484-8
; Sequence 8, Application US/10216484
; Publication No. US20030103976A1
; GENERAL INFORMATION:
; APPLICANT: Serizawa, No. US20030103976A1ufusa
; APPLICANT: Haruyama, Hideyuki
; APPLICANT: Nakahara, Kaori
; APPLICANT: Tamaki, Ikuko
; APPLICANT: Takahashi, Tohru
; TITLE OF INVENTION: Anti-Fas Antibodies
; FILE REFERENCE: 980126CIP/HG
; CURRENT APPLICATION NUMBER: US/10/216,484
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US/09/499,662
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: US 09/053,583
; PRIOR FILING DATE: 1998-04-01
; NUMBER OF SEQ ID NOS: 165
; SEQ ID NO 8
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1392)
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (58)..(1392)
```



[illegible][illegible]





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 30, 2003, 22:11:21 ; Search time 1841.13 Seconds  
(without alignments)  
2019.725 Million cell updates/sec

Title: US-08-836-455-4  
Perfect score: 816  
Sequence: 1 MECSWFLFLSITGTGHSQ.....TVSSAKTTPPVYPLVPGSL 153

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US08836455/runat\_29082003\_132902\_22290/app\_query.fasta\_1.654  
-DB=EST -QFMT=fastap -SUFFIX=pn.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -START=ext -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US08836455.ecgn\_1.1.3596.e/runat\_29082003\_132902\_22290 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estm: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hic: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hic: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: em\_gss\_hum: \*  
18: em\_gss\_inv: \*  
19: em\_gss\_pln: \*  
20: em\_gss\_vrt: \*  
21: em\_gss\_fun: \*  
22: em\_gss\_mam: \*  
23: em\_gss\_mus: \*  
24: em\_gss\_pro: \*  
25: em\_gss\_rod: \*  
26: em\_gss\_phq: \*  
27: em\_gss\_vrl: \*  
28: gb\_gss1: \*

29: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	603	73.9	739	10	BE284158	BE284158 601099428
2	600	73.5	637	14	BY724721	BY724721 BY724721
3	596	72.0	644	10	BE281961	BE281961 601099518
4	589.5	72.2	1576	11	AK007918	AK007918 Mus muscu
5	588.5	72.1	767	10	BF165456	BF165456 60177347
6	585	71.7	892	10	EG518664	EG518664 602578528
7	578.5	70.9	941	10	BF138189	BF138189 601781314
8	577.5	70.8	794	12	BI150371	BI150371 602915205
9	575	70.5	771	10	BF140551	BF140551 601787584
10	570.5	69.9	529	14	CA576389	CA576389 K0644A12-
11	570.5	69.9	546	14	CA576594	CA576594 K0647B06-
12	570	69.9	896	13	BU524012	BU524012 AGENCOURT
13	567	69.5	847	10	BF165486	BF165486 601777393
14	562.5	68.9	616	10	BF136295	BF136295 601781413
15	561.5	68.8	667	14	BY724790	BY724790 BY724790
16	559	68.5	885	10	BF162852	BF162852 601767129
17	549	67.3	540	14	CA579750	CA579750 K0741H07-
18	548.5	67.2	507	10	BF015548	BF015548 uy23a08.y
19	548.5	67.2	793	10	BF136093	BF136093 601783539
20	548	67.2	519	14	CA580087	CA580087 K0746G01-
21	547.5	67.1	529	14	CA577954	CA577954 K0716F08-
22	547.5	67.1	891	10	BF138460	BF138460 601782916
23	547	67.0	748	10	BF136397	BF136397 601783927
24	544.5	66.7	958	13	BU523796	BU523796 AGENCOURT
25	544	66.7	700	10	BF540088	BF540088 602050325
26	543.5	66.6	538	14	CA578968	CA578968 K0731A04-
27	543.5	66.6	604	10	BF140035	BF140035 601791175
28	543.5	66.6	724	12	BG962137	BG962137 602826902
29	542.5	66.5	469	14	CA574907	CA574907 K0622H01-
30	541	66.3	477	14	CA570382	CA570382 K0505G11-
31	540	66.2	745	12	BG871607	BG871607 602790090
32	539	66.1	1012	10	BF579009	BF579009 602096127
33	538.5	66.0	598	12	BI104341	BI104341 602889919
34	537.5	65.9	685	10	BF136104	BF136104 601783550
35	537	65.8	464	10	BE368025	BE368025 601221914
36	536.5	65.7	429	14	CA579143	CA579143 K0733D01-
37	536.5	65.7	475	14	CA578534	CA578534 K0725B11-
38	536.5	65.7	1003	14	CB236612	CB236612 AGENCOURT
39	535.5	65.6	516	14	CA579084	CA579084 K0732F03-
40	535	65.6	724	10	BF168514	BF168514 601775412
41	534.5	65.5	783	10	BF143757	BF143757 601789643
42	534.5	65.5	896	12	BI149320	BI149320 602848859
43	531.5	65.1	933	13	BQ929209	BQ929209 AGENCOURT
44	530.5	65.0	829	10	BF144014	BF144014 601791270
45	528.5	64.8	465	14	CA578116	CA578116 K0718G11-

ALIGNMENTS

RESULT 1  
BE284158  
LOCUS BE284158  
DEFINITION 601099428F1 NCI\_CGAP\_Lu29 Mus musculus cdna clone IMAGE:3491766 5',  
mRNA linear EST 13-JUL-2000  
mRNA sequence.  
ACCESSION BE284158  
VERSION BE284158.1  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 739)



cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## FEATURES

## source

```

1. .637
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A53001123"
/sex="male"
/tissue_type="aorta and vein"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male aorta and vein"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAATAATTAATCCGCCGCCGCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
```

BASE COUNT 147 a 179 c 160 g 151 t

## ORIGIN

## Alignment Scores:

```

Pred. No.: 1.29e-52 Length: 637
Score: 600.00 Matches: 116
Percent Similarity: 83.55% Conservative: 11
Best Local Similarity: 76.32% Mismatches: 25
Query Match: 73.53% Indels: 0
DB: 14 Gaps: 0
```

US-08-836-455-4 (1-153) x BY724721 (1-637)

```

Qy 1 MetGluCysSerTrpValPheLeuPheLeuSerIleThrGlyValHisSerGln 20
Db 52 ATGGAATGGAGCAGAGCTTTATCTTCTCTATCATGAAGTGCAGGTGTCTACTCCAG 111
Qy 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 112 GTCCAGCTGGCAGAGCTGGAGCTGAGTGGTGAAGCCTGGAGTTCAGTGAAGGTGCC 171
Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstPvAllyGlnThrPro 60
Db 172 TCGAAGCTTCTGGATACGCCTTCACCTAATTTACTGTAGTGGGTAAAGCAGAGCCCT 231
Qy 61 GlyClnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80
Db 232 GGACAGGCGCTTGGTGGATTGGAGTGATTAATCTCTGGAAGTGTGTACTAATACTAAT 291
Qy 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerThrAlaTyrMet 100
Db 292 GAGAAATTCAGGCGAGGCAACACTGACTGCGACAAATCTCCACGACTGGCTCATATG 351
Qy 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120
Db 352 CAGCTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATTCTTGTGCAAGATTCGCC 411
```

```

Qy 121 GLuGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThr 140
Db 412 CGAGGGGAGTGTGACTACTGGGCGCAAGCAGCAGCTCTCAGAGTCTCTCAGCCAAACG 471
Qy 141 ThrProProValTyrProLeuValProGlySer 152
Db 472 ACACCCCATCTGCTCTATCAGCTGGCCCTGGATCT 507

RESULT 3
BE281961 644 bp mRNA linear EST 26-OCT-2000
LOCUS 601099518F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3492028 5',
DEFINITION mRNA sequence.
ACCESSION BE281961
VERSION BE281961.1 GI:9157200
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 644)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8537 row: c column: 05
High quality sequence stop: 561.
FEATURES
Location/Qualifiers
1. .644
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3492028"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu29"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 147 a 176 c 155 g 166 t
ORIGIN

Alignment Scores:
Pred. No.: 3.42e-52 Length: 644
Score: 596.00 Matches: 115
Percent Similarity: 80.12% Conservative: 14
Best Local Similarity: 71.43% Mismatches: 20
Query Match: 73.04% Indels: 12
DB: 10 Gaps: 2

US-08-836-455-4 (1-153) x BE281961 (1-644)
Qy 1 MetGluCysSerTrpValPheLeuPheLeuSerIleThrGlyValHisSerGln 20
Db 6 ATGGAATGGAGCTGGGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 65
Qy 21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 66 GTTCAGCTGCAACAGTCTGAGCTGAGTGTGGTGACCTGGAGCTTCAAGTATATCC 125
Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstPvAllyGlnThrPro 60
Db 411 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstPvAllyGlnThrPro 60
```





```

||||||| ||| ||||||||||| |||
455 AACACAGCCCATCGGTCTATCCACTGGCCCT 488

Bg518664          892 bp      mRNA      linear      EST 02-APR-2001
602578528F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3491766 5',
mRNA sequence.
Bg518664
Bg518664.1 GI:13513763
EST.
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 892)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM8536 row: h column: 07
High quality sequence stop: 771.
Location/Qualifiers
FEATURES
source
1.892
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:3491766"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/clone_lib="NCI_CGAP_Lu29"
/lab_host="DH10B"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      217 a      266 c      211 g      198 t
ORIGIN
Alignment Scores:
Pred. No.:      7,43e-51      Length:      892
Score:      585.00      Matches:      113
Percent Similarity:      82.35%      Conservative:      13
Best Local Similarity:      73.86%      Mismatches:      26
Query Match:      71.69%      Indels:      1
DB:      10      Gaps:      0

US-08-836-455-4 (1-153) x Bg518664 (1-892)

Qy      1 MetGluCysSerTrpValPheLeuLeuSerIleThrGlyValHisSerGln 20
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
33 ATGGAATGGATCTGGATCTTCTCTCATCTCTCAGGAAGTGCAGTCCACG 92

Qy      21 AlaTyrLeuGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLys-MetSe 40
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
93 GTTCAGCTGCAGCAGTCTGGAGCTGAGCTGGCGAGCGCTGGGCTTCAGTGAACGATGTC 152

Qy      40 rCysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMetHisTrpValLysGlnThrPr 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
153 CTGCAAGGCTTCGGTTACTCATTCACAGCTATGATAGCTGGGTGAAGCAGAAAC 212

Qy      60 oGlyGlnGlyLeuLurPrIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTras 80
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
213 TAGACAGGCGCTTGAGTGGATTGGAGAGATTATCTCTGGAAGTGGTAACACTTACTACAA 272

```

---

```

Qy      80 nGlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrIleTyrMe 100
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
273 TGAAGAAGTTCAAGGCAAGGCCACTGCTGCTGAGACAAATCTCCAGCACACGCTACAT 332

Qy      100 tGlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTr 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
333 GCAGCTCAGCAGCTGACATCTGAGGACTCTGCAGTCTATTTCTGTGCAAGATGGGGGA 392

Qy      120 pGluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysTh 140
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
393 GCCTCGGAGCTGGACTACTGGGCGCAAGCCACCATCATCATCTCTCAGCCAAAC 452

Qy      140 rThrProProValTyrProLeuValProGlySer 152
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
453 GACACCCCATCTGTCTATCCACTGGCCCTGGATCT 489

RESULT 7
Bg518189
LOCUS
DEFINITION
601781314F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4009467 5',
mRNA sequence.
Bg518189
Bg518189.1 GI:10977229
VERSION
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 941)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM9245 row: k column: 04
High quality sequence stop: 637.
Location/Qualifiers
FEATURES
source
1.941
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4009467"
/tissue_type="tumor, metastatic to mammary"
/clone_lib="NCI_CGAP_Lu30"
/lab_host="DH10B"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; transgenic model WNT-1, expression driven by
MTMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dt. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT      232 a      266 c      248 g      195 t
ORIGIN
Alignment Scores:
Pred. No.:      3.81e-50      Length:      941
Score:      578.50      Matches:      112
Percent Similarity:      85.14%      Conservative:      14
Best Local Similarity:      75.68%      Mismatches:      19
Query Match:      70.89%      Indels:      3
DB:      10      Gaps:      1

US-08-836-455-4 (1-153) x Bg518189 (1-941)

Qy      6 ValPheLeuPheLeuLeuSerIleThrGlyValHisSerGlnAlaTyrLeuGlnGln 25

```





```

source      1. 771
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6Ncr"
/db_xref="taxon:10090"
/clone="IMAGE:4015172"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI-CCAP_Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; transgenic model WNT-1, expression driven by
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 179 a 215 c 195 g 182 t
ORIGIN
Alignment Scores:
Pred. No.: 6.8e-50 Length: 771
Score: 575.00 Matches: 110
Percent Similarity: 82.00% Conservatave: 13
Best Local Similarity: 73.33% Mismatches: 25
Query Match: 70.47% Indels: 2
DB: 10 Gaps: 1
US-08-836-455-4 (1-153) x BFL40551 (1-771)
Qy 1 MetGluCysSerTrpValPheLeuLeuSerIleThrThrGlyValHisSerGln 20
Db 15 ATGGAGTGGGCTGGCTTCTCTCTCTGTCAGTAACTGAGGTGTCACACGAG 74
Qy 21 AlaTyrLeuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 75 GTTCAGTGAAGCAGCTCTGGAGCTGAGCTGGTGAAGCTGGCTGAGTGAAGCTGCC 134
Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMetHisTrpValLysGlnThrPro 60
Db 135 TCGAAGGCTTCTGCTACCTTCAACAGCTATATATGCTGGTGAAGCAGAGGCT 194
Qy 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80
Db 195 GGACAGGGCCCTGAGTGGTAGGATACCTTTCTCTTACATGATGATACATAATGCAAT 254
Qy 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100
Db 255 GAGAAGTTCAAAGGCAAGGCCACACTGAGGACTCTGAGGACTCTGCTTACTTACTGCAAGGGGTACTAC 314
Qy 101 GlnIleSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120
Db 315 GAGCTCAGCAGCCTGACCTCTGAGGACTCTGCGCTCTATTACTTGTGCAAGGGGTACTAC 374
Qy 121 GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThr 140
Db 375 -----GTATTGTACTACTGGGGCCAAAGGCACCACTATCATCAGTCTCTCTCAGCCAAACA 428
Qy 141 ThrProProValTyrProLeuValPro 150
Db 429 ACAGCCCATCGGCTATATCCACTGGCCCT 458
RESULT 10
CA576389 529 bp mRNA linear EST 19-NOV-2002
LOCUS K0644A12-5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)
DEFINITION IMAGE:30072971 5', mRNA sequence.
CA576389
ACCESSION CA576389.1 GI:25121091
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 529)
AUTHORS Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Carter,M.G.,
Taub,D., Longo,D.L., Keller,J. and Ko,M.S.H.
TITLE Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Jin-/c-Kit-/Sca-1+) cDNA Library (Long)
JOURNAL Unpublished
COMMENT Other_Ests: K0644A12-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0644 row: A column: 12
Seq primer: M13 Reverse
High quality sequence stop: 529
POLYA=No.
FEATURES
Location/Qualifiers
1..529
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6Ncr"
/db_xref="niaEST:K0644A12-5N"
/db_xref="taxon:10090"
/clone="NIA:K0644A12 IMAGE:30072971"
/tissue_type="Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)"
/dev_stage="Age approx.10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1+) cDNA Library (Long)"
/notes="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an Oligo(dT) primer [invitrogen:
5'-pGACTAGTCTAGTCGCGCGCCGCTTTTCTTTT-3'] from
1.1 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lona-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT1 plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.2 kb. The library was constructed
by Yulan Piao (NIA)."
BASE COUNT 125 a 138 c 136 g 130 t
ORIGIN
Alignment Scores:
Pred. No.: 1.22e-49 Length: 529
Score: 570.50 Matches: 110
Percent Similarity: 84.11% Conservatave: 17
Best Local Similarity: 72.85% Mismatches: 24
Query Match: 69.91% Indels: 2
DB: 14 Gaps: 0
US-08-836-455-4 (1-153) x CA576389 (1-529)
Qy 1 MetGluCysSerTrpValPheLeuLeuSerIleThrThrGlyValHisSerGln 20
Db 25 ATGGAATGGACCTGGGTCTTCTCTCTCTGTCAGTAACTGAGGTGTCACCTCCAG 84
Qy 21 AlaTyrLeuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 85 GTTCAGCTGCAGCAGCTCTGGAGCTGAGCTGATGAAGCTGGGCGCTCAGTGAAGCTTCC 144

```







[illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 31, 2003, 00:02:06 ; Search time 3130.85 Seconds  
(without alignments)  
1999.191 Million cell updates/sec

Title: US-08-836-455-4  
Perfect score: 153  
Sequence: 1 MECSWFLFLLSITGVHQS.....TVSSAKTTPPPVPLVPGSL 153

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Word size: 1  
Total number of hits satisfying chosen parameters: 5773148  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US08836455/runat\_29082003\_132953\_23045/app\_query.fasta\_1.654  
-DB=GenEmbl -QEXT=fastap -SUFFIX=p2noli.rge -MINMATCH=0\_1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFM=ptp  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US08836455\_@CGN\_1\_1\_3608\_@runat\_29082003\_132953\_23045 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pin.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	153	100.0	461	6	AR164506	AR164506 Sequence
2	153	100.0	461	6	BD085738	BD085738 Methods o
3	135	88.2	411	10	AF124720	AF124720 Mus muscu
4	38	24.8	471	10	AB050080	AB050080 Mus muscu
5	31	20.3	458	6	AR083801	AR083801 Sequence
6	31	20.3	458	6	AR198720	AR198720 Sequence
7	31	20.3	458	6	AR275320	AR275320 Sequence
8	29	19.0	294	10	AF303872	AF303872 Mus muscu
9	27	17.6	735	12	SC0250760	AJ250760 Mus muscu
10	22	14.4	357	10	MMIG07	X58580 Mouse hybri
11	22	14.4	357	10	MMIG09	M54977 Mus muscu
12	22	14.4	484	10	MUSIGH4C11	M24785 Mouse anti-
13	22	14.4	490	10	MUSIGH4C11A	AF433159 Mus muscu
14	21	13.7	390	10	AF433159	M97876 Mouse hybri
15	21	13.7	390	10	MUSL77IGHV	M13330 Mouse Ig ga
16	21	13.7	631	10	MUSIGHEB	U90237 Mus muscu
17	21	13.7	902	12	XMU90237	U49832 Synthetic s
18	21	13.7	279	10	MMIGGCVRG	Z25453 M.musculus
19	20	13.1	282	10	MMIGGCVRB	Z25443 M.musculus
20	20	13.1	285	10	MMIGGCVRC	Z25451 M.musculus
21	20	13.1	290	10	MMIGGCVRF	AF303873 Mus muscu
22	20	13.1	294	10	AF303873	AY172396 M.musculus
23	20	13.1	297	10	AY172396	Z25457 M.musculus
24	20	13.1	302	10	MMIGGCVRI	Z25447 M.musculus
25	20	13.1	303	10	MMIGGCVRD	Z25449 M.musculus
26	20	13.1	303	10	MMIGGCVRE	AY171955 Mus muscu
27	20	13.1	306	10	AY171955	AF328128 Mus muscu
28	20	13.1	309	10	AF328128	AY171952 Mus muscu
29	20	13.1	321	10	AY171952	AB043346 Mus muscu
30	20	13.1	333	10	AB043346	Z73339 M.musculus
31	20	13.1	352	10	MMMD50C	Z73342 M.musculus
32	20	13.1	355	10	MMMD47C	AY247151 Mus muscu
33	20	13.1	356	10	AY247151	AF178586 Mus muscu
34	20	13.1	357	10	AF178586	Z73357 M.musculus
35	20	13.1	358	10	MMMD01C	Z73338 M.musculus
36	20	13.1	358	10	MMMD52C	AY229957 Mus muscu
37	20	13.1	360	10	AY229957	AF277091 Synthetic
38	20	13.1	360	12	AF277091	U16182 Mus muscu
39	20	13.1	366	10	MMU16182	AY058907 Mus muscu
40	20	13.1	387	10	AY058907	AB069863 Mus muscu
41	20	13.1	405	10	AB069863	U24114 Mus muscu
42	20	13.1	453	10	MMU24114	AF178830 Mus muscu
43	20	13.1	457	10	AY178830	M19901 Mouse Ig re
44	20	13.1	469	10	MMSIGHAAE	
45	20	13.1	469	10	MMSIGHAAE	

ALIGNMENTS

AR164506 LOCUS AR164506 461 bp DNA linear PAT 17-OCT-2001  
 DEFINITION Sequence 3 from patent US 6274143.  
 ACCESSION AR164506  
 VERSION AR164506.1 GI:16237556  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 461)  
 AUTHORS Chatterjee, M. and Foon, K.A.  
 TITLE Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11D10  
 JOURNAL Patent: JP 2001523269-A 3 14-AUG-2001;  
 FEATURES Location/Qualifiers  
 source 1..461  
 BASE COUNT 112 a 121 c 119 g 109 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,98e-155 Length: 461  
 Score: 153.00 Matches: 153  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-08-836-455-4 (1-153) x AR164506 (1-461)

Qy 1 MetGluCysSerTrpValPheLeuPheLeuSerIleThrThrGlyValHisSerGln 20  
 |||||  
 Db 1 ATGGAATGCAGCTGGGTCTTCTCTCTCTGTCATTAACACAGGTGCCACTCCAG 60  
 |||||  
 Qy 21 AlaTyrLeuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
 |||||  
 Db 61 GCATTATCTACAGCAGTCTGGGGCTGAGCTGGTGGCTCAGTGAAGATGTC 120  
 |||||  
 Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstTrpValLysGlnThrPro 60  
 |||||  
 Db 121 TGCAGAGCTTCTGGCTACACATTGACCAATGACATGACATGCGGTAAAGCAGACACCT 180  
 |||||  
 Qy 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80  
 |||||  
 Db 181 GGACAGGCTGGGAATGGATTGGAAATATTTTCTCTGAAATGGTGATCTACTACAAAT 240  
 |||||  
 Qy 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerThrAlaTyrMet 100  
 |||||  
 Db 241 CAGAAATTTAAGGCAAGGCTCATTTGACTGCAGACACATCTCTCCAGCAGACGCTACATG 300  
 |||||  
 Qy 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120  
 |||||  
 Db 301 CAGATCAGCAGCCTGACATCTGAAGACTCTGGGTCTATTCTCTGCAAGGGAAGTGG 360  
 |||||  
 Qy 121 GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThr 140  
 |||||  
 Db 361 GAGGGTCTCTGGACTACTGGGTCAAGGAACCTCAGTCACCGCTCTCTCCAGCCAAAACG 420  
 |||||  
 Qy 141 ThrProProValTyrProLeuValProGlySerLeu 153  
 |||||  
 Db 421 ACACCCCAACCCGCTATATCCACTGGTCCCTGGAAAGCTTG 459  
 |||||

RESULT 2  
 BD085738 LOCUS BD085738 461 bp DNA linear PAT 27-AUG-2002  
 DEFINITION Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11D10.  
 ACCESSION BD085738  
 VERSION BD085738.1 GI:22631348  
 KEYWORDS JP 2001523269-A/2.  
 SOURCE unidentified  
 ORGANISM unclassified.

REFERENCE 1 (bases 1 to 461)  
 AUTHORS Chatterjee, M. and Foon, K.A.  
 TITLE Methods of delaying development of HMFG-associated tumors using anti-idiotypic antibody 11D10  
 JOURNAL Patent: JP 2001523269-A 2 20-NOV-2001;  
 COMMENT THE UNIVERSITY OF KENTUCKY RESEARCH FOUNDATION  
 OS Unidentified  
 PN JP 2001523269-A/2  
 PD 20-NOV-2001  
 PF 12-JUN-1998 JP 1999503252  
 PR 13-JUN-1997 US 60/049540, 11-JUN-1998 US 09/096244 P1  
 PC A61K39/395, A61K39/39//C07K16/42  
 CC A61K39/395, A61K39/39//C07K16/42  
 CC Strandedness: Single;  
 CC Topology: Linear;  
 CC Methods of delaying development of HMFG-associated tumors CC using  
 CC anti-idiotypic antibody 11D10  
 FH Key Location/Qualifiers  
 FT CDS 1..459  
 FT mat peptide 58.  
 FEATURES Location/Qualifiers  
 source 1..461  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"  
 BASE COUNT 112 a 121 c 119 g 109 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,98e-155 Length: 461  
 Score: 153.00 Matches: 153  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-08-836-455-4 (1-153) x BD085738 (1-461)

Qy 1 MetGluCysSerTrpValPheLeuPheLeuSerIleThrThrGlyValHisSerGln 20  
 |||||  
 Db 1 ATGGAATGCAGCTGGGTCTTCTCTCTCTGTCATTAACACAGGTGCCACTCCAG 60  
 |||||  
 Qy 21 AlaTyrLeuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40  
 |||||  
 Db 61 GCATTATCTACAGCAGTCTGGGGCTGAGCTGGTGGCTCAGTGAAGATGTC 120  
 |||||  
 Qy 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstTrpValLysGlnThrPro 60  
 |||||  
 Db 121 TGCAGAGCTTCTGGCTACACATTGACCAATGACATGACATGCGGTAAAGCAGACACCT 180  
 |||||  
 Qy 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80  
 |||||  
 Db 181 GGACAGGCTGGGAATGGATTGGAAATATTTTCTCTGAAATGGTGATCTACTACAAAT 240  
 |||||  
 Qy 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerThrAlaTyrMet 100  
 |||||  
 Db 241 CAGAAATTTAAGGCAAGGCTCATTTGACTGCAGACACATCTCTCCAGCAGACGCTACATG 300  
 |||||  
 Qy 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp 120  
 |||||  
 Db 301 CAGATCAGCAGCCTGACATCTGAAGACTCTGGGTCTATTCTCTGCAAGGGAAGTGG 360  
 |||||  
 Qy 121 GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThr 140  
 |||||  
 Db 361 GAGGGTCTCTGGACTACTGGGTCAAGGAACCTCAGTCACCGCTCTCTCCAGCCAAAACG 420  
 |||||  
 Qy 141 ThrProProValTyrProLeuValProGlySerLeu 153  
 |||||  
 Db 421 ACACCCCAACCCGCTATATCCACTGGTCCCTGGAAAGCTTG 459  
 |||||

RESULT 3  
 AF124720

AF124720	411 bp	mRNA	linear	ROD 22-MAY-2000
Mus musculus immunoglobulin heavy chain mRNA, partial cds.				
AF124720				
AF124720.1	GI:14164544			
Mus musculus (house mouse)				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
1 (bases 1 to 411)				
Tripathi,P.K., Qin,H., Bhattacharya-Chatterjee,M., Ceriani,R.L., Foon,K.A. and Chatterjee,S.K.				
Construction and characterization of a chimeric fusion protein consisting of an anti-idiotypic antibody mimicking a breast cancer-associated antigen and the cytokine GM-CSF				
Hybridoma 18 (2), 193-202 (1999)				
99306687				
PUBMED				
10380019				
2 (bases 1 to 411)				
Chatterjee,S.K. and Tripathi,P.K.				
Direct Submission				
Submitted (29-JAN-1999)				
Internal Medicine, University of Kentucky,				
800 Rose Street, Lexington, KY 40536, USA				
Location/Qualifiers				
1..411				
/organism="Mus musculus"				
/mol_type="mRNA"				
/strain="BALB/c"				
/db_xref="taxon:10090"				
1..>411				
/note="anti-idiotypic antibody 11b10; mimics a breast cancer-associated antigen, human fat globule (HMFG)"				
/codon_start=1				
/product="immunoglobulin heavy chain"				
/protein_id="AAK55119.1"				
/db_xref="GI:14164545"				
/translation="MCSWVFLLESLITGVHVSQAVLQQSGAELVRSAGSVKMSCRASGYLTSTNMHWYKOTPGQLIEWIFNPPGNGDYINOKFKGRASLTADTSSSYAYMOI SUTTSDESAVFYCARGNWEGALYDQGGSVTYSS"				
BASE COUNT	102 a	101 c	107 g	101 t
ORIGIN				
Alignment Scores:				
Pred. No.:	8.56e-136	Length:	411	
Score:	135.00	Matches:	135	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	88.24%	Indels:	0	
DB:	10	Gaps:	0	
US-08-836-455-4 (1-153) x AF124720 (1-411)				
Qy	3	CysSerTrpValPheLeuPheLeuSerIleThrThrGlyValHisSerGlnAlaTyr	22	
Dd	7	TGCAGCTGGGTCTTTCTTCCTCCCTGTCATAACTACAGTGCTCCACTCCCAGGCTTAT	66	
Qy	23	LeuGlnGlnSerGlyAlaGluLeuValArgSerGlyValAsrValllysMetSerCysLys	42	
Dd	67	CTACACAGCTCTGGGCGCTGNAGCTGGTAGGTCTGGGCCCTCACTGAAGATGCTCCTGCAG	136	
Qy	43	AlaSerGlyTyrThrLeuThrSerTyrAsnMetHisTrpValLysGlnThrProGlyGln	62	
Dd	127	GCYTCTGGCTACACATTGACCAGCTTACAATATGCACCTGGGTAAAGCAGACACCTGGCAG	186	
Qy	63	GlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrAsnGlnLys	82	
Dd	187	GGCTCGGAATGGATTCGAATATTTTTCTCGAAATGGTGATCTACTACAAATCAGAAG	246	
Qy	83	PheLysGlyLysAlaSerLeuThrAlaAspThr-SerSerThrAlaTyrMetGlnIle	102	
Dd	247	TTTAAGGGCAGGCCCTCATTCAGCTGCAGACACATCTCCAGCACAGCCTACATGCAGATC	306	
Qy	103	SerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrpGluGly	122	

Db 307 AGCAGCCTGACATCTGAAGACTCTGGCGTCTATTCTGTCAAGAGGAACTGGGAGGT 366  
 Qy 123 AlaleuAspTyrTrpGlyGlnGlyThrSerValThrValSerSer 137  
 Db 367 GCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGCTCTCCTCA 411

**RESULT 4**  
 AB050080  
 LOCUS Mus musculus VH9H8 mRNA for anti-A/U antibody, partial cds.  
 DEFINITION  
 AB050080  
 VERSION  
 AB050080.1 GI:19909935  
 KEYWORDS  
 Mus musculus (house mouse)  
 SOURCE  
 Mus musculus  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**REFERENCE**  
 1 Kitagawa,Y. and Okuhara,E.  
 TITLE The separation of three antibody populations from  
 anti-poly(A).poly(U) antibodies elicited in mice or rabbits and  
 antigenic features of poly(A).poly(U)  
 JOURNAL Mol. Immunol. 19 (2), 257-266 (1982)  
 MEDLINE 82245325  
 PUBMED 6178956  
**REFERENCE**  
 2 Kitagawa,Y.  
 TITLE anti-dsRNA (A/U) Ab VH region VH9H8  
 JOURNAL Published Only in Database (2002)  
 REFERENCE 3 (bases 1 to 471)  
 AUTHORS Kitagawa,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-OCT-2000) Yoshichika Kitagawa, Akita Prefectural  
 University, Biotechnology Institute; Minami 2-2, Ogata, Akita  
 010-0444, Japan (E-mail:kitagawa@agri.akita-pu.ac.jp,  
 URL:www.akita-pu.ac.jp/ Tel:81-185-45-2026(ex.400),  
 Fax:81-185-45-2678)

**FEATURES**  
 source  
 1. .471  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 /cell\_line="hybridoma 9H8"  
 gene  
 1. .471  
 /gene="VH9H8"  
 11. .>469  
 /gene="VH9H8"  
 /codon\_start=1  
 /product="anti-A/U antibody"  
 /protein\_id="BAB87192.1"  
 /db\_xref="GI:19909936"  
 /translation="MGWTGVFLILSTTTGVHSGAELVRSGASVKMSCKAS  
 GYFTSYNMHWKOTPGQGLIEYIFLNAGTNYNOKFKGKATLTADTSSSTAYMIOIS  
 SLTSDSAVYFCARGEYGNPWFAYWGQGLIVTWSAAKTTAQPVPYPLAPOS"  
 CDS  
 114 a 126 c 121 g 110 t  
 BASE COUNT  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2,65e-31 Length: 471  
 Score: 38.00 Matches: 38  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 24.84% Indels: 0  
 DB: 10 Gaps: 0

US-08-836-455-4 (1-153) x AB050080 (1-471)

Qy 10 LeuLeuSerIleThrThrGlyValHisSerGlnAlaTyrLeuGlnSerGlyValaGlu 29  
 Db 38 CTCCTGTCATACTACAGTGTCACCTCCAGGCTTATCTACAGCAGTCTGGGCTGAA 97

QY 30 LeuValArgSerGlyAlaSerValLysMetSerCysLysAlaSerGlyTyrThr 47  
|||||  
Db 98 CTGTGAGGCTCTGGGGCTCAGTGAAGATCTCTGCAAGGCTTCTGGCTACACA 151  
RESULT 5  
AR083801  
LOCUS AR083801 458 bp DNA linear PAT 01-SEP-2000  
DEFINITION Sequence 3 from patent US 5977316.  
ACCESSION AR083801  
VERSION AR083801.1 GI:10010572  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 458)  
AUTHORS Chatterjee,M., Foon,K.A. and Chatterjee,S.K.  
TITLE Monoclonal antibody 1A7 and related polypeptides  
JOURNAL Patent: US 5977316-A 3 02-NOV-1999;  
FEATURES Location/Qualifiers  
source  
BASE COUNT 106 a 131 c 114 g 107 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.89e-24 Length: 458  
Score: 31.00 Matches: 31  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.26% Indels: 0  
DB: 6 Gaps: 0  
US-08-836-455-4 (1-153) x AR083801 (1-458)  
QY 123 AlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrPro 142  
|||||  
Db 364 GCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCAAAAGACACCC 423  
QY 143 ProProValTyrProLeuValProGlySerLeu 153  
|||||  
Db 424 CCACCCGCTATCCATTGGTCCCTGGAAGCTTG 456  
RESULT 6  
AR198720  
LOCUS AR198720 458 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 3 from patent US 6355244.  
ACCESSION AR198720  
VERSION AR198720.1 GI:20248794  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 458)  
AUTHORS Foon,K.A. and Chatterjee,M.  
TITLE Methods and compositions for the treatment of psoriasis  
JOURNAL Patent: US 6355244-A 3 12-MAR-2002;  
FEATURES Location/Qualifiers  
source  
BASE COUNT 106 a 131 c 114 g 107 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.89e-24 Length: 458  
Score: 31.00 Matches: 31  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.26% Indels: 0  
DB: 6 Gaps: 0  
US-08-836-455-4 (1-153) x AR198720 (1-458)  
QY 123 AlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrPro 142

|||||  
Db 364 GCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCAAAAGACACCC 423  
QY 143 ProProValTyrProLeuValProGlySerLeu 153  
|||||  
Db 424 CCACCCGCTATCCATTGGTCCCTGGAAGCTTG 456  
RESULT 7  
AR275320  
LOCUS AR275320 458 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 3 from patent US-6509016.  
ACCESSION AR275320  
VERSION AR275320.1 GI:29708411  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 458)  
AUTHORS Chatterjee,M., Foon,K.A. and Chatterjee,S.K.  
TITLE Monoclonal antibody 1A7 and use for the treatment of melanoma and small cell carcinoma  
JOURNAL Patent: US 6509016-A 3 21-JAN-2003;  
FEATURES Location/Qualifiers  
source  
BASE COUNT 106 a 131 c 114 g 107 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.89e-24 Length: 458  
Score: 31.00 Matches: 31  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.26% Indels: 0  
DB: 6 Gaps: 0  
US-08-836-455-4 (1-153) x AR275320 (1-458)  
QY 123 AlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrPro 142  
|||||  
Db 364 GCTCTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCAGCAAAAGACACCC 423  
QY 143 ProProValTyrProLeuValProGlySerLeu 153  
|||||  
Db 424 CCACCCGCTATCCATTGGTCCCTGGAAGCTTG 456  
RESULT 8  
AF303872  
LOCUS AF303872 294 bp mRNA linear ROD 25-SEP-2001  
DEFINITION Mus musculus clone J558.41 Immunoglobulin heavy chain variable region mRNA, partial cds.  
ACCESSION AF303872  
VERSION AF303872.1 GI:11612050  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 294)  
AUTHORS Haines,B.B., Angeles,C.V., Parmelee,A.P., McLean,P.A. and Brodeur,P.H.  
TITLE Germline diversity of the expressed BALB/c VHJ558 gene family  
JOURNAL Mol. Immunol. 38 (1), 9-18 (2001)  
MEDLINE 21376477  
PUBMED 11483206  
REFERENCE 2 (bases 1 to 294)  
AUTHORS Haines,B.B., Angeles,C.V., Parmelee,A.P., McLean,P.A. and Brodeur,P.H.  
TITLE Direct Submission  
JOURNAL Submitted (08-SEP-2000) Pathology, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111, USA  
FEATURES Location/Qualifiers  
source  
I. .294





```

/db_xref="taxon:10090"
/cell_line="Hybridoma 12S18-G"
1. .357
/gene="IG heavy chain"
/mRNA
<1. .>357
/gene="IG heavy chain"
/note="variable region"
/evidence-experimental
<1. .>357
/gene="IG heavy chain"
/note="variable region"
/codon_start=1
/product="IG heavy chain"
/protein_id="CAA41456.1"
/db_xref="GI:930150"
/translation="QVQLQEPGAEVVKPGASVKMSCKASGYTFTNYMHVYKQPGQG
LEMGNIPGSDYTNNEKPKSKATLVDTSSSYAYMQISLTSSEDSAVYYCARNEG
AWYFDVNGXGTTVAXSS"
V_region
<1. .>357
/gene="IG heavy chain"
/product="IG heavy chain"
/note="variable region"
BASE COUNT      88 a      87 c      95 g      78 t      9 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      3.44e-14      Length:      357
Score:          22.00      Matches:      22
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      14.38%      Indels:      0
DB:              10      Gaps:      0

```

US-08-836-455-4 (1-153) x MMIG07 (1-357)

```

Qy  92 AspThrSerSerThrAlaTyMetGlnIleSerSerLeuThrSerGluAspSerAla 111
|||||
Db  217 GACACATCTCCAGACAGCCATGATGAGTACGACGCCGTGACATCTGAGACTCTGCG 276
|||||
Qy  112 ValTyr 113
|||||
Db  277 GTCTAT 282

```

```

RESULT 11
MMIG09
LOCUS      Mouse hybridoma 12S18 mRNA for immunoglobulin heavy chain V region.
DEFINITION X58582 Y00794
ACCESSION  X58582.1 GI:51593
VERSION    Ig heavy chain; Ig variable region; immunoglobulin.
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Meek, K., Hasemann, C., Pollok, B., Alkan, S.S., Brait, M., Slaoui, M.,
Urban, J., and Capra, J.D.
AUTHORS    Structural characterization of antiidiotypic antibodies. Evidence
TITLE      that Ab2s are derived from the germline differently than Ab1s
JOURNAL    J. Exp. Med. 169 (2), 519-533 (1989)
MEDLINE    89094248
PUBMED     2492056
REFERENCE  2 (bases 1 to 357)
AUTHORS    Meek, K.
TITLE      Direct Submission
JOURNAL    Submitted (09-JAN-1989) K. Meek, University of Texas Southwestern
75235, USA

```

```

FEATURES
source
1. .357
/organism="Mus musculus"
/mol_type="mRNA"
/strain="A/J"

```

```

/db_xref="taxon:10090"
/cell_line="Hybridoma 12S28-G"
1. .357
/gene="IG heavy chain"
/mRNA
<1. .>357
/gene="IG heavy chain"
/note="variable region"
/evidence-experimental
<1. .>357
/gene="IG heavy chain"
/note="variable region"
/codon_start=1
/product="IG heavy chain"
/protein_id="CAA41458.1"
/db_xref="GI:930152"
/translation="QVQLQEPGAEVVKPGASVKMSCKASGYTFTNYMHVYKQPGQG
LEMGNIPGSDYTNNEKPKSKATLVDTSSSYAYMQISLTSSEDSAVYYCARNEG
AWYFDVNGXGTTVAXSS"
V_region
<1. .>357
/gene="IG heavy chain"
/product="IG heavy chain"
/note="variable region"
BASE COUNT      89 a      86 c      96 g      80 t      6 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      3.44e-14      Length:      357
Score:          22.00      Matches:      22
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      14.38%      Indels:      0
DB:              10      Gaps:      0

```

US-08-836-455-4 (1-153) x MMIG09 (1-357)

```

Qy  92 AspThrSerSerThrAlaTyMetGlnIleSerSerLeuThrSerGluAspSerAla 111
|||||
Db  217 GACACATCTCCAGACAGCCATGATGAGTACGACGCCGTGACATCTGAGACTCTGCG 276
|||||
Qy  112 ValTyr 113
|||||
Db  277 GTCTAT 282

```

```

RESULT 12
MUSIGH4C11
LOCUS      Mus musculus immunoglobulin heavy chain precursor (IGH) mRNA,
DEFINITION partial cds.
ACCESSION  M54977
VERSION    M54977.1 GI:194528
KEYWORDS   C-region; V-region; immunoglobulin heavy chain.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 484)
AUTHORS    Sood, A.K., Cheng, H.D. and Kohler, H.
TITLE      An efficient and general method for sequencing immunoglobulin mRNAs
JOURNAL    J. Immunol. Methods 95 (2), 227-235 (1986)
MEDLINE    87084812
PUBMED     2432131
COMMENT    Original source text: Mouse anti-idiotypic hybridoma cell line
4C11, cDNA to mRNA.
FEATURES
source
1. .484
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/map="chromosome 12"
/cell_line="4C11"
/tissue_type="anti-idiotypic hybridoma"
1. .484
/gene="Igh"

```

```

CDS
37..>484
/gene="IGH"
/codon_start=1
/product="immunoglobulin heavy chain precursor"
/protein_id="AA51051.1"
/db_xref="GI:553945"
/translation="MGWSWIFLELLSGTAVLSEVLOQSGPELLKPGASVKISCRAS
GYFTDTHHWKQSKGLEWIGGDPNDYNTFYNEKFKDKATLVDRSSSTAYMEL
RSLTSDSAVYCYASYDYGALDYWGQGTSTVTSSTKTPPSVYPLAP"
37..93
/gene="IGH"
/mat_peptide
94..>484
/gene="IGH"
/product="immunoglobulin heavy chain"
445..>484
/gene="IGH"
BASE COUNT 120 a 134 c 112 g 118 t
ORIGIN
Alignment Scores:
Pred. No.: 4.59e-14 Length: 484
Score: 22.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.38% Indels: 0
DB: 10 Gaps: 0
US-08-836-455-4 (1-153) x MUSIGH4C11 (1-484)
QY 122 GYAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerAlaLysThrThr 141
Db 397 GGTGCTCTGGACTACTGGGTCAAGAACCTCAGTACCCTCTCCTCAGCCAAACGACA 456
QY 142 ProPro 143
Db 457 CCCCCA 462
RESULT 13
MUSIG4C11A
LOCUS MUSIG4C11A 490 bp mRNA linear ROD 27-APR-1993
DEFINITION Mouse anti-Idotype immunoglobulin heavy chain variable, constant,
and complementarity determining regions 1-3 (4C11) mRNA.
ACCESSION M24785.1 GI:194133
VERSION M24785.1
KEYWORDS C-region; V-region; anti-Idotypic antibody; complementarity
determining region; immunoglobulin heavy chain.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 490)
AUTHORS Cheng,H.L., Sood,A.K., Ward,R.E., Kieber-Emmons,T. and Kohler,H.
TITLE Structural basis of stimulatory anti-Idotypic antibodies
JOURNAL Mol. Immunol. 25 (1), 33-40 (1988)
MEDLINE 88142863
PUBMED 3125424
COMMENT Original source text: Mus musculus (strain BALB/c, sub_species
domesticus) cDNA to mRNA.
FEATURES
source
Location/Qualifiers
1..490
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/sub_species="domesticus"
/db_xref="taxon:10090"
1..490
/gene="4C11"
V_region 1..441
/gene="4C11"
184..198
/gene="CDR1"
misc_feature 184..198
/gene="CDR1"

CDS
37..>484
/gene="IGH"
/codon_start=1
/product="immunoglobulin heavy chain precursor"
/protein_id="AA51051.1"
/db_xref="GI:553945"
/translation="MGWSWIFLELLSGTAVLSEVLOQSGPELLKPGASVKISCRAS
GYFTDTHHWKQSKGLEWIGGDPNDYNTFYNEKFKDKATLVDRSSSTAYMEL
RSLTSDSAVYCYASYDYGALDYWGQGTSTVTSSTKTPPSVYPLAP"
37..93
/gene="IGH"
/mat_peptide
94..>484
/gene="IGH"
/product="immunoglobulin heavy chain"
445..>484
/gene="IGH"
BASE COUNT 120 a 134 c 116 g 120 t
ORIGIN
Alignment Scores:
Pred. No.: 4.64e-14 Length: 490
Score: 22.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.38% Indels: 0
DB: 10 Gaps: 0
US-08-836-455-4 (1-153) x MUSIG4C11A (1-490)
QY 122 GYAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerAlaLysThrThr 141
Db 397 GGTGCTCTGGACTACTGGGTCAAGAACCTCAGTACCCTCTCCTCAGCCAAACGACA 456
QY 142 ProPro 143
Db 457 CCCCCA 462
RESULT 14
AF433159
LOCUS AF433159 390 bp mRNA linear ROD 20-NOV-2001
DEFINITION Mus musculus monoclonal antibody K1-11 heavy chain variable region
mRNA, partial cds.
ACCESSION AF433159
VERSION AF433159.1 GI:17016945
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 390)
AUTHORS Ho,M. and Segre,M.
TITLE Mimicry of cocaine by anti-Idotypic antibodies
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 390)
AUTHORS Ho,M. and Segre,M.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2001) Pathobiology, University of Illinois at
Urbana-Champaign, 2001 South Lincoln Avenue, Urbana, IL 61802, USA
FEATURES
source
Location/Qualifiers
1..390
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="1A:D5:A5:C5"
/sex="female"
/cell_type="hybridoma"
/tissue_type="hyperimmunized spleen"
/dev_stage="adult"
<1..>390
/note="anti-Idotypic antibody; K1-11 VH"
/codon_start=1
/product="monoclonal antibody K1-11 heavy chain variable
region"
/protein_id="AAL33541.1"
/db_xref="GI:17016946"

```

Search completed: August 31, 2003, 03:58:14  
Job time : 3132.85 secs

QY 123 AlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaIysThrThrPro 142  
 316 GCGTTGGACTACTGGGGTCAAGAACTCAGTCACCGTCTCCTCAGCCAAAACGACACCC 375  
 Db

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 30, 2003, 22:12:36 ; Search time 230.013 Seconds  
(without alignments)  
1795.608 Million cell updates/sec

Title: US-08-836-455-4  
Perfect score: 153  
Sequence: 1 MECSWFLFLSITGVHSQ.....TVSSAKTTPPPYPLVPGSL 153

Scoring table:  
OLIGO 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delopt 6.0 , Delopt 7.0

Searched: 2552756 seqs, 1349719017. residues

Word size: 1

Total number of hits satisfying chosen parameters: 5103490

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV=xlh  
-MODEL=frame+ p2n.model -DB=US08836455/runat\_29082003.132953.23037/app\_query.fasta\_1.654  
-O=/cn2\_1/USPTO\_spool/US08836455/runat\_29082003.132953.23037/app\_query.fasta\_1.654  
-DB=N\_Geneseq 15Jun03 -QFMT=fastap -SUFFIX=p2noli.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human-0.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US08836455@cgn\_1\_1\_401\_@runat\_29082003.132953.23037 -NCPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N\_Geneseq.19Jun03.\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	153	100.0	461	18	AAT85150	Murine monoclonal
2	153	100.0	461	20	AAV83773	Antibody 1LD10 hea
3	153	100.0	461	25	AA151274	Mouse 1LD10 antibo
4	31	20.3	458	17	AAT31333	Anti-idiotypic mono
5	31	20.3	458	20	AAZ31366	MAB 1A7 heavy chain
6	31	20.3	458	20	AAZ89553	Heavy chain variab
7	30	19.6	458	20	AAZ60630	Monoclonal antibod
8	27	17.6	366	20	AAZ82028	Anti-STX1 heavy ch
9	22	14.4	861	16	AAQ81500	sfv anti-rev seque
10	22	14.4	861	16	AAZ45347	Single chain sfv a
11	20	13.1	765	22	AAZ85590	DNA encoding a fus
12	20	13.1	765	22	AAZ85591	DNA encoding a fus
13	20	13.1	1239	22	AAZ85591	DNA encoding a fus
14	20	13.1	1280	22	AAZ85564	Anti-CD20 single c
15	20	13.1	1925	21	AAZ15019	DNA encoding a CD-
16	19	12.4	360	18	AAZ96345	CDNA for Ig heavy
17	19	12.4	373	15	AAZ71395	Anti-carcinoembryo
18	19	12.4	373	15	AAZ71396	Anti-carcinoembryo
19	19	12.4	399	16	AAZ90425	DNA encoding anti-
20	19	12.4	402	16	AAZ37471	Sequence encoding
21	19	12.4	402	16	AAZ90426	DNA encoding anti-
22	19	12.4	402	17	AAZ47894	MAB NM-01 heavy ch
23	19	12.4	420	22	AAZ81910	Anti-CA125 bifunct
24	19	12.4	447	20	AAZ57786	Anti-HCV Ser/Thr p
25	19	12.4	458	10	AAZ91146	2H7 Vh sequence.
26	19	12.4	459	20	AAZ82357	Mouse antibody 2H7
27	19	12.4	459	22	AAZ22070	2H7 heavy chain va
28	19	12.4	470	16	AAZ99892	Mouse VIA-4 antibo
29	19	12.4	470	18	AAZ74760	Alpha-4 integrin m
30	19	12.4	474	18	AAZ70808	Mouse anti-idiotyp
31	19	12.4	474	18	AAZ70810	Mouse anti-idiotyp
32	19	12.4	477	18	AAZ70806	Mouse anti-idiotyp
33	19	12.4	486	18	AAZ59339	MH1 monoclonal ant
34	19	12.4	491	18	AAZ70868	2H1 heavy chain va
35	19	12.4	491	19	AAZ18557	Mouse 2H7 antibody
36	19	12.4	491	19	AAZ03926	Mouse 2H7 antibody
37	19	12.4	491	19	AAZ18593	Mouse 2H7 antibody
38	19	12.4	518	18	AAZ36316	2H7 antibody heavy
39	19	12.4	520	18	AAZ51042	Coding sequence fo
40	19	12.4	588	14	AAZ43385	H-chain V-region o
41	19	12.4	626	25	ACC44926	TSH receptor antib
42	19	12.4	626	25	ACC44927	TSH receptor antib
43	19	12.4	636	17	AAZ87815	Antibody 7G12 heav
44	19	12.4	642	25	ACC44914	TSH receptor antib
45	19	12.4	642	25	ACC44915	TSH receptor antib

ALIGNMENTS

RESULT 1  
AAT85150  
ID AAT85150 standard; CDNA; 461 BP.  
XX AAT85150;  
AC AAT85150;  
XX  
DT 25-MAR-2003 (updated)  
DT 04-JAN-1998 (first entry)  
XX  
DE Murine monoclonal anti-idiotypic antibody 1LD10 VH CDNA.  
XX  
KW Monoclonal antibody 1LD10; anti-idiotypic antibody; mucin;  
KW human milk fat globule; HMF; tumour; breast cancer; vaccine; ss.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers

Db	241	CAGAAGTTTAAAGGCAAGGCGCTCATTTGACTCAGACACATCTCTCCAGCACACGCTTACATG	300
Qy	101	GlnIleSerSerIeuThrSerGluAspSerAlaValTyrPheCysAlaArgGlyAsnTrp	120
Db	301	CAGATCAGCAGCCCTGACATCTGAAGACTCTCGGGTCTATTCTGTGCAGAGGGGAACTCGG	360
Qy	121	GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThr	140
Db	361	GAGGGTGCTCTGGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCTCAGCCAAAACG	420
Qy	141	ThrProProValTyrProLeuValProGlySerIeu	153
Db	421	ACACCCCAACCCGCTATCCACTGGTCCCTCGAAGCTTG	459
RESULT 2			
ID	AAV83773	standard; cDNA; 461 BP.	
XX	AAV83773;		
XX	AC		
XX	16-MAR-1999	(first entry)	
DE	Antibody 1LD10	heavy chain variable region coding sequence.	
KW	Murine; mouse;	antibody; light chain; variable region; anti-idiotypic; ss;	
KW	human milk fat globule;	tumour; ovary; lung; pancreas; carcinoma; breast.	
OS	Mus sp.		
FH	Key	Location/Qualifiers	
FT	CDS	1..459	
FT	FT	/*tag= a	
FT	FT	/product= "antibody 1LD10 heavy chain variable region"	
FT	FT	/note= "no stop codon is given at the 3' end of the	
FT	FT	sequence"	
FT	sig_peptide	1..57	
FT	FT	/*tag= b	
FT	mat_peptide	58..459	
FT	FT	/*tag= c	
XX			
PN	W09856419-A1.		
XX			
PD	17-DEC-1998.		
XX			
PF	12-JUN-1998;	98WO-US12250.	
XX			
PR	11-JUN-1998;	98US-0096244.	
PR	13-JUN-1997;	97US-0049540.	
XX			
PA	(RENT ) UNIV KENTUCKY RES FOUND.		
XX			
PI	Chatterjee M, Foon KA;		
XX			
DR	WPI; 1999-060029/05.		
DR	P-PSDB; AAW87594.		
XX			
PT	Delaying development of, or treating, HMFG-associated tumours -		
PT	using anti-idiotypic antibody 1LD10 raised against antibodies to		
PT	human milk fat globule protein		
XX			
PS	Disclosure; Fig 2; 54pp; English.		
XX			
CC	This sequence represents the coding sequence for the murine antibody		
CC	1LD10 heavy chain variable region. This anti-idiotypic antibody is used		
CC	to delay the development of, or treat, a human milk fat globule (HMFG)		
CC	associated tumour in an individual having low tumour burden.		
CC	The antibody 1LD10 is used to prevent the recurrence of HMFG-associated		
CC	tumours e.g. ovarian, non-small cell lung and pancreatic carcinoma,		
CC	especially for treating breast tumours.		
XX			
SQ	Sequence 461 BP; 112 A; 121 C; 119 G; 109 T; 0 other;		
Alignment Scores:			



ID XX AAT31333 standard; cDNA; 458 BP.  
AC AAT31333;  
XX 25-MAR-2003 (updated)  
XX 26-FEB-1997 (first entry)  
XX  
XX Anti-idiotype monoclonal antibody 1A7 variable heavy chain, cDNA.  
XX Murine; mouse; anti-idiotype; monoclonal antibody; MAb; 1A7;  
XX variable heavy chain; ganglioside 2; GD2; 14G2a; neuroblastoma;  
XX glycosphingolipid; human; neuroectodermal; tumour; glioma; lung;  
XX malignant melanoma; soft tissue sarcoma; small cell carcinoma;  
XX vaccine; treatment; palliate; detection; diagnosis;  
XX recombinant production; purification; probe; primer; assay;  
XX amplification; gene therapy; ss.  
XX  
XX Mus musculus.  
XX  
XX Key Location/Qualifiers  
XX mat\_peptide 1..456  
XX /\*tag- a  
XX /\*transl\_except- pos:373..375, aa:Trp  
XX  
XX W09622373-A2.  
XX  
XX 25-JUL-1996.  
XX  
XX 17-JAN-1996; 96WO-US00882.  
XX  
XX 17-JAN-1995; 95US-0372676.  
XX 16-JAN-1996; 96US-0591196.  
XX 17-JAN-1995; 95US-0372676.  
XX 16-JAN-1996; 96US-0591196.  
XX  
XX (KENT ) UNIV KENTUCKY.  
XX  
XX Chatterjee M, Chatterjee SK, Foon KA;  
XX WPI; 1996-354530/35.  
XX P-PSDB; AAW03200.  
XX  
XX Monoclonal antibody 1A7 and related polynucleotide(s) and  
XX polypeptide(s) - useful to treat or palliate a GD2-associated  
XX disease, e.g. melanoma and glioma  
XX  
XX Claim 11; Fig 2; 141pp; English.  
XX  
XX The present sequence encodes the murine anti-idiotype monoclonal  
XX antibody (MAb) 1A7 variable heavy chain. MAb 1A7 was raised against  
XX the anti-ganglioside 2 (GD2) MAb 14G2a, which binds an unique  
XX epitope of GD2. As the glycosphingolipid GD2 is expressed at high  
XX density by human neuroectodermal tumours, e.g. malignant melanoma,  
XX neuroblastoma, glioma, soft tissue sarcoma and small cell carcinoma  
XX of the lung, MAb 1A7, or its cDNA can be used in a vaccine to treat  
XX or palliate such diseases. They can also be used to reduce the  
XX risk of recurrence of a clinically detectable tumour, and detect an  
XX anti-GD2 Ab bound to a tumour cell.  
XX MAb 1A7 overcomes immune tolerance and induces an immune response  
XX against GD2, which comprises anti-GD2 Ab (humoral response) and  
XX GD2-specific cells (cellular response). It can be used to purify  
XX anti-1A7 (AB3), anti-GD2 (Abl') or 14G2a (Abl), detect anti-1A7 or  
XX anti-GD2 in a sample or measure the level of cellular anti-1A7 or  
XX anti-GD2 activity.  
XX The cDNA can be used in expression systems for 1A7 prodn., and in  
XX the prepn. of probes and primers to respectively assay for 1A7  
XX cDNA, and amplify desired polynucleotides for use in gene therapy.  
XX (Updated on 25-MAR-2003 to correct PR field.)  
XX  
XX SQ Sequence 458 BP; 106 A; 131 C; 114 G; 107 T; 0 other;  
XX  
XX Alignment Scores: 7.77e-20 Length: 458  
XX Pred. No.: 458

Score: 31.00 Matches: 31  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.26% Indels: 0  
DB: 17 Gaps: 0  
US-08-836-455-4 (1-153) x AAT31333 (1-458)  
QY 123 AlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrPro 142  
DB 364 GCTCTGACTACTGGGTCAAGGAACCTCAGTCACCGTCTCTCTCAGCAAAACGACACC 423  
QY 143 ProProValTyrProLeuValProGlySerLeu 153  
DB 424 CCACCGTCTATCCATTGGTCTCCTGGAAGCTTG 456  
RESULT 5  
AAZ31366  
ID AAZ31366 standard; cDNA; 458 BP.  
XX  
XX AAZ31366;  
XX  
XX 07-FEB-2000 (first entry)  
XX  
XX MAb 1A7 heavy chain variable region encoding cDNA.  
XX  
XX Monoclonal antibody; MAb; 1A7; GD2; immune response; melanoma;  
XX neuroblastoma; glioma; soft tissue carcinoma; small cell carcinoma;  
XX tumor-associated antigen; ss.  
XX  
XX Synthetic.  
XX OS Mus sp.  
XX  
XX US5977316-A.  
XX  
XX 02-NOV-1999.  
XX  
XX 16-JAN-1996; 96US-0591196.  
XX  
XX 17-JAN-1995; 95US-0372676.  
XX  
XX (KENT ) UNIV KENTUCKY.  
XX  
XX Foon KA, Chatterjee SK, Chatterjee M;  
XX WPI; 1999-619711/53.  
XX P-PSDB; AAY49210.  
XX  
XX Monoclonal antibody 1A7 which elicits an anti-GD2 immunological  
XX response, useful for the development of products for the detection and  
XX treatment of cancers -  
XX Disclosure; Fig 2; 74pp; English.  
XX  
XX The invention provides a monoclonal antibody (MAb) designated 1A7, which  
XX elicits an anti-GD2 (tumor-associated antigen) immunological response in  
XX humans. MAb 1A7 has defined light and heavy chain variable region  
XX sequences. The MAb 1A7 and polypeptides can be used for eliciting an  
XX anti-GD2 immune response. The polypeptides can also be used for detecting  
XX or purifying anti-GD2 antibody. The products can be used for treating GD2  
XX -associated diseases, e.g. melanoma, neuroblastoma, glioma, soft tissue  
XX carcinoma, and small cell carcinoma. They can be used for palliating the  
XX disease or for reducing the risk of recurrence. The present sequence  
XX represents the cDNA encoding the heavy chain variable region of MAb 1A7.  
XX  
XX SQ Sequence 458 BP; 106 A; 131 C; 114 G; 107 T; 0 other;  
XX  
XX Alignment Scores: 7.77e-20 Length: 458  
XX Pred. No.: 31 Matches: 31  
XX Score: 31.00 Mismatches: 0  
XX Percent Similarity: 100.00% Conservative: 0  
XX Best Local Similarity: 100.00% Mismatches: 0  
XX Query Match: 20.26% Indels: 0



DB: 20 Gaps: 0

US-08-836-455-4 (1-153) x AAZ31366 (1-458)

QY 123 AlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrThrPro 142  
 |||||  
 Db 364 GCTCTGGACTACTGGGGTCAAGAACTCAGTCACCGTCTCTCTCAGCCAAAGACACCC 423

QY 143 ProProValTyrProLeuValProGlySerLeu 153  
 |||||  
 Db 424 CCACCGTCTATCATTTGTCCTCGAAGCTTG 456

RESULT 6  
 ID AAX89553 standard; cDNA; 458 BP.  
 XX  
 AC AAX89553;  
 XX  
 DT 06-OCT-1999 (first entry)  
 XX  
 DE Heavy chain variable region of MAb 1A7.  
 XX  
 DE heavy chain variable region; antibody 1A7; T cell response; melanoma;  
 KW ganglioside GD2; CDR; complementarity determining region; carcinoma; ss.  
 KW  
 XX  
 OS Mus musculus.  
 XX  
 EH Key Location/Qualifiers  
 FT CDS 1..456  
 FT /\*tag= a  
 FT /product= "vh chain MAb 1A7  
 FT /note= "No stop codon given"  
 FT sig\_peptide 1..57  
 FT /\*tag= b  
 FT mat\_peptide 58..456  
 FT /\*tag= c  
 XX  
 PN US5935821-A.  
 XX  
 PD 10-AUG-1999.  
 XX  
 PF 21-NOV-1996; 96US-0752844.  
 XX  
 PR 21-NOV-1996; 96US-0752844.  
 PR 17-JAN-1995; 95US-0372676.  
 PR 16-JAN-1996; 96US-0591196.  
 XX  
 PA (KENT ) UNIV KENTUCKY.  
 XX  
 PI Chatterjee M, Chatterjee SK, Foon KA;  
 XX  
 DR WPI; 1999-457600/38.  
 DR P-PSDB; AAV28469.  
 XX  
 PT Anti-GD2 immunological peptides useful for the treatment of tumours  
 PT especially melanomas and small cell carcinomas  
 XX  
 PS Claim 7: Fig 2; 84pp; English.  
 XX  
 CC The sequence is the variable heavy chain region of monoclonal  
 CC anti-idiotypic antibody 1A7. The polypeptide encoded by this sequence has  
 CC three CDRs (complementarity determining regions). When administered to  
 CC an individual the 1A7 antibody induces an immune response against  
 CC ganglioside GD2. The light chain variable region of the 1A7 antibody  
 CC (AAV28469) is also capable of eliciting an anti GD2 response in mammals.  
 CC Both the heavy and light chain variable regions of the 1A7 antibody  
 CC produce anti-GD2 T cell and antibody responses. The peptides and  
 CC antibodies may be useful for the modulation of ganglioside GD2, and  
 CC particularly for the treatment of GD2-associated tumours (e.g. melanoma,  
 CC neuroblastoma, glioma, soft tissue sarcoma, and small cell carcinoma  
 CC (including small cell lung cancer)).  
 XX  
 SQ Sequence 458 BP; 106 A; 131 C; 114 G; 107 T; 0 other;

Alignment Scores:  
 Pred. No.: 7.77e-20 Length: 458  
 Score: 31.00 Matches: 31  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 20.26% Indels: 0  
 DB: 20 Gaps: 0

US-08-836-455-4 (1-153) x AAX89553 (1-458)

QY 123 AlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrThrPro 142  
 |||||  
 Db 364 GCTCTGGACTACTGGGGTCAAGAACTCAGTCACCGTCTCTCTCAGCCAAAGACACCC 423

QY 143 ProProValTyrProLeuValProGlySerLeu 153  
 |||||  
 Db 424 CCACCGTCTATCATTTGTCCTCGAAGCTTG 456

RESULT 7  
 AAX60630  
 ID AAX60630 standard; cDNA; 458 BP.  
 XX  
 AC AAX60630;  
 XX  
 DT 03-AUG-1999 (first entry)  
 XX  
 DE Monoclonal antibody 1A7 heavy chain variable region encoding cDNA.  
 XX  
 DE Psoriasis; immunological response; anti-idiotypic antibody; glutamate;  
 KW chronic plaque; pustular; plaque-type psoriasis; psoriatic arthritis;  
 KW monoclonal antibody; 1A7; ss.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9925380-A2.  
 XX  
 PD 27-MAY-1999.  
 XX  
 PF 17-NOV-1998; 98WO-US24607.  
 XX  
 PR 16-NOV-1998; 98US-0192838.  
 PR 17-NOV-1997; 97US-0065774.  
 XX  
 PA (KENT ) UNIV KENTUCKY RES FOUND.  
 XX  
 PI Chatterjee M, Foon KA;  
 XX  
 DR WPI; 1999-347407/29.  
 DR P-PSDB; AAY21546.  
 XX  
 PT Treatment of psoriasis  
 XX  
 PS Disclosure; Fig 3; 48pp; English.  
 XX  
 CC The invention provides a method of treating of psoriasis by administering  
 CC an antigen which has similar immunogenic properties to an antigen  
 CC expressed on cells of psoriatic tissue so that an immunological response  
 CC is elicited in the individual. The antigen stimulates the generation of  
 CC anti-idiotypic antibodies that neutralize the aberrant immune response  
 CC causing the psoriasis. The method is used to treat psoriasis, especially  
 CC chronic plaque, glutamate, pustular, plaque-type psoriasis or psoriatic  
 CC arthritis. The compositions allow the individual's own immune system to  
 CC act against psoriatic tissue. The present sequence represents a cDNA  
 CC encoding the heavy chain variable region of monoclonal antibody 1A7.  
 XX  
 SQ Sequence 458 BP; 106 A; 132 C; 113 G; 107 T; 0 other;

Alignment Scores:  
 Pred. No.: 6.67e-19 Length: 458  
 Score: 30.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 19.61% Indels: 0  
 DB: 20 Gaps: 0

US-08-836-455-4 (1-153) x AAX60630 (1-458)

Qy 123 AlalaLeuaspYrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrThrPro 142  
 Db 364 GCTCTGGACTACTCGGTCAGGAACTCAGTCACCTCTCTCTCAGCAAAACGACACCC 423  
 Qy 143 ProProValTyrProLeuValProGlySer 152  
 Db 424 CCACCCGTCATCCATTGGTCCCTGGGAGC 453

RESULT 8  
 AAX82028  
 ID AAX82028 standard; DNA; 366 BP.  
 XX  
 AC AAX82028;  
 XX  
 DT 10-SEP-1999 (first entry)  
 XX  
 DE Anti-STX1 heavy chain variable region encoding DNA.  
 XX  
 KW Humanised; monoclonal antibody; Mab; Shiga toxin; immunoglobulin;  
 KW Ig constant region; enterohemorrhagic Escherichia coli; EHEC; uremia;  
 KW edema; bloody diarrhoea; hemorrhagic colitis; hemolytic uremic syndrome;  
 KW thrombocytopenia; EHEC-mediated disease; anti-STX1; ss.  
 XX  
 OS Shigella dysenteriae.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..366  
 FT /\*tag= a  
 FT /transl\_except= (pos;67..60, aa:Asp)  
 FT /note= "the start and stop codons are not indicated"

WO9932645-A1.  
 XX  
 PD 01-JUL-1999.  
 XX  
 PF 22-DEC-1998; 98WO-US27267.  
 XX  
 PR 18-DEC-1998; 98US-0215163.  
 PR 23-DEC-1997; 97US-0086635.  
 XX  
 PA (MELT/) MELTON-CELSA A.  
 PA (OBRI/) O'BRIEN A D.  
 PA (SCHM/) SCHMITT C K.  
 PA (STIN/) STINSON J L.  
 PA (WONG/) WONG H.  
 XX  
 PI Melton-Celsa A, O'Brien AD, Schmitt CK, Stinson JL;  
 PI Wong H;  
 XX  
 DR WPI; 1999-418935/35.  
 DR P-PSDB; AAY21816.  
 XX  
 PT Humanized monoclonal antibodies against Shiga toxins, useful for  
 PT protection against enterohemorrhagic Escherichia coli or other Shiga  
 PT toxin producing bacteria  
 XX  
 PS Claim 11; Fig 3; 75pp; English.  
 XX  
 CC The invention relates to humanised monoclonal antibodies (MAB) against  
 CC Shiga toxins. The humanised MAB that binds to Shiga toxin comprises a  
 CC constant and a variable region, where: (a) the constant region contains  
 CC at least part of a human immunoglobulin (Ig) constant region; and (b) the  
 CC variable region contains at least part of a non-human Ig variable region.  
 CC Host cells transformed with vectors encoding a humanised MAB against  
 CC Shiga toxin type 2 is useful for treating a patient with an infection  
 CC caused by enterohemorrhagic Escherichia coli (EHEC) or other Shiga toxin  
 CC producing bacteria. The humanised MAB can also be used to reduce illness  
 CC caused by EHEC or other Shiga toxin producing bacteria. EHEC are

CC associated with food-borne outbreaks of bloody diarrhoea (hemorrhagic  
 CC colitis) and the hemolytic uremic syndrome. In particular, the humanised  
 CC MABs ameliorate edema, thrombocytopenia and uremia associated with EHEC-  
 CC mediated disease. The present sequence represents a DNA encoding an anti  
 CC -STX1 heavy chain variable region.

XX  
 SQ Sequence 366 BP; 94 A; 92 C; 92 G; 88 T; 0 other;

Alignment Scores:  
 Pred. No.: 3.42e-16 Length: 366  
 Score: 27.00 Matches: 27  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 17.65% Indels: 0  
 DB: 20 Gaps: 0

US-08-836-455-4 (1-153) x AAX82028 (1-366)

Qy 89 LeuThrAlaAspThrSerSerSerThrAlaTyrMetGlnIleSerSerLeuThrSerGlu 108  
 Db 208 TTGACTCGACACATCTCTCCAGCAGCCTCATCATGATCATGACATCTGACATCTGAA 267  
 Qy 109 AspSerAlaValTyrPheCys 115  
 Db 268 GACTCTCGGTCTATTCTGT 288

RESULT 9  
 AAO81500  
 ID AAO81500 standard; cDNA; 861 BP.  
 XX  
 AC AAO81500;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 30-AUG-1995 (first entry)  
 XX  
 DE sfv anti-rev sequence.  
 XX  
 KW HIV-1; human immunodeficiency virus type 1; AIDS; Rev protein;  
 KW intracellular immunization; gene therapy; single chain antibody; Fv;  
 KW sfv; antibody engineering; resistance; cell immunity; Hela; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9503832-A1.  
 XX  
 PD 09-FEB-1995.  
 XX  
 PF 28-JUL-1994; 94WO-US08448.  
 XX  
 PR 30-JUL-1993; 93US-0099870.  
 XX  
 PA (UYJE-) UNIV JEFFERSON THOMAS.  
 XX  
 PI Duan L, Pomerantz R;  
 XX  
 DR WPI; 1995-082039/11.  
 XX  
 PT Method for conducting gene therapy - comprises using recombinant  
 PT gene encoding antibody binding antigen associated with a disease;  
 PT useful for providing cell immunity.  
 XX  
 PS Example 4; Page 23; 62pp; English.  
 XX  
 CC The sequence given in AAO81500 encodes an sfv anti-rev antibody  
 CC consisting of the variable domains of the heavy and light chains  
 CC of a mouse MAB against HIV-1 IIIB rev. The sfv specifically  
 CC binds a highly conserved Rev domain. HeLa T4 cells expressing the  
 CC sfv were resistant to all HIV-1 clinical isolates tested.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 861 BP; 199 A; 234 C; 230 G; 198 T; 0 other;

Alignment Scores:

Pred. No.: 3.53e-11 Length: 861  
Score: 22.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 14.38% Indels: 0  
DB: 16 Gaps: 0

US-08-836-455-4 (1-153) x AAQ81500 (1-861)

QY 126 TyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrProProVal 145  
|||||  
DB 794 TACTGGGTCAGGAACCTCAGTCACCGCTCTCTCAGCCAAACGACACCCCGTC 853  
QY 146 TyrPro 147  
|||||  
DB 854 TATCCT 859

## RESULT 10

AA45347

ID AA45347 standard; cDNA; 861 BP.

XX AC AA45347;

XX DT 08-AUG-1997 (first entry)

XX DE Single chain sfv anti-rev antibody encoding cDNA.

XX KW Gene therapy; antibody; immunisation; human immunodeficiency virus;

XX KW HIV; human T-cell leukaemia virus; ss.

XX OS Mus musculus.

XX PN WO637234-AL.

XX PD 28-NOV-1996.

XX PF 23-MAY-1996; 96WO-US07393.

XX PR 23-MAY-1995; 95US-0447610.

XX PA (UYJE-) UNIV JEFFERSON THOMAS.

XX PI Duan L, Pomerantz RJ;

XX DR WPI; 1997-020948/02.

XX Improved gene therapy using recombinant gene coding for an antibody

PT - for intracellular immunisation against pathogens recognised by the

PT antibody, esp. human immunodeficiency virus HIV-1

XX Example 4; Page 54; 213pp; English.

XX The present sequence encodes a single chain sfv anti-rev antibody  
CC constructed using variable domains of the heavy and light chains of a  
CC murine monoclonal antibody against (HIV-1/IIIB) rev (the parent  
CC antibody). This is incorporated into a viral vector where  
CC expression of the anti-rev gene causes inhibition of the rev function  
CC and so affects replication of the other virus (HIV). Rev is one of the  
CC essential regulatory proteins of HIV, it binds to rev responsive element  
CC (RRE) and promotes the nuclear export, stabilisation and utilisation of  
CC the viral mRNA's containing RRE. A novel gene therapy method has been  
CC produced, where a recombinant (rev) gene is introduced into the cells of  
CC a mammal. The method is improved by using a rec gene encoding an  
CC antibody (Ab) that is selectively specific for an intracellular (IC)  
CC antigen associated with a disease. The method is used to prevent or  
CC halt the progress of a disease by IC immunisation. Specifically, the Ab  
CC can be used to inhibit the replication of a virus, such as human T-cell  
CC leukaemia virus or especially HIV-1, or of other pathogens, e.g.  
CC bacteria, fungi. The method provides immunity before or after the  
CC development of the disease and can be used to control the severity of  
CC the disease.

XX Sequence 861 BP; 199 A; 234 C; 230 G; 198 T; 0 other;

## Alignment Scores:

Pred. No.: 3.53e-11 Length: 861  
Score: 22.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 14.38% Indels: 0  
DB: 18 Gaps: 0

US-08-836-455-4 (1-153) x AA45347 (1-861)

QY 126 TyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrProProVal 145  
|||||  
DB 794 TACTGGGTCAGGAACCTCAGTCACCGCTCTCTCAGCCAAACGACACCCCGTC 853

QY 146 TyrPro 147  
|||||

DB 854 TATCCT 859

## RESULT 11

AA45347

ID AA45347 standard; DNA; 765 BP.

XX AC AA45347;

XX DT 02-APR-2001 (first entry)

XX DE DNA encoding a fusion of a single chain antibody and streptavidin.

XX KW Streptavidin; tumour cell; cancer; adenocarcinoma;

XX KW hematological malignancy; ss.

XX OS Synthetic.

XX OS Streptomyces avidinii.

XX OS Unidentified.

XX PN WO200075333-AL.

XX PD 14-DEC-2000.

XX PF 05-JUN-2000; 2000WO-US15595.

XX PR 07-JUN-1999; 99US-0137900.

XX PR 03-DEC-1999; 99US-0168976.

XX PA (NEOR-) NEORX CORP.

XX PI Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;

XX DR WPI; 2001-091213/10.

XX New vector constructs for expressing genomic streptavidin fusion  
PT proteins which are useful for targeting tumour cells associated with  
PT cancer, e.g. adenocarcinomas -  
XX Example 5; Page 95; 100pp; English.

XX The present sequence encodes a fusion of an anti-CD20 single chain  
CC antibody and streptavidin. The fusion protein is expressed using  
CC vectors of the invention. The specification describes vector constructs  
CC for expressing streptavidin fusion proteins. The vector comprises a  
CC nucleic acid encoding streptavidin or its functional variant operatively  
CC linked to a promoter, and a cloning site for insertion of a second  
CC nucleic acid sequence encoding a polypeptide to be fused with  
CC streptavidin, interposed between the promoter and the first nucleic  
CC acid sequence. Alternatively, the vector construct comprises a nucleic  
CC acid, operatively linked to a promoter, encoding a polypeptide to be  
CC fused with streptavidin, and a cloning site for insertion of a second  
CC nucleic acid encoding at least 129 amino acids of streptavidin or its  
CC functional variant. The fusion proteins are useful for targeting tumour  
CC cells, particularly tumour cells associated with cancer.  
CC e.g. adenocarcinomas or hematological malignancies. The vector construct  
CC is useful for expressing of streptavidin fusion proteins. In particular,

CC these are useful as tools for medical diagnostics and therapeutic  
CC purposes, e.g. for detecting the presence or absence of, or treating, a  
CC target site within a mammalian host.

XX  
SQ Sequence 765 BP; 169 A; 201 C; 230 G; 165 T; 0 other;

Alignment Scores:  
Pred. No.: 2.33e-09 Length: 765  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.07% Indels: 0  
DB: 22 Gaps: 0

US-08-836-455-4 (1-153) x AAC86590 (1-765)

QY 49 ThrSerTyrAsnMetHisTrpValLysGlnThrProGlyGlnGlyLeuGluTrpIleGly 68  
Db 88 ACCAGTTACAAATATGCACTGGGTAAACGACACCTGGACAGGCCCTGGGATGGATTGGA 147

RESULT 12

AAC86591  
ID AAC86591 standard; DNA; 765 BP.

XX AAC86591;

XX 02-APR-2001 (first entry)

XX DNA encoding a fusion of a single chain antibody and streptavidin.

XX Streptavidin; tumour cell; cancer; adenocarcinoma;  
KW hematological malignancy; ss.

XX Synthetic.

OS Streptomyces avidinii.

OS Unidentified.

XX WO200075333-A1.

XX 14-DEC-2000.

XX 05-JUN-2000; 2000WO-US15595.

XX 07-JUN-1999; 99US-0137900.

XX 03-DEC-1999; 99US-0168976.

XX (NEOR-) NEORX CORP.

XX Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;

XX WPI; 2001-091213/10.

XX New vector constructs for expressing genomic streptavidin fusion  
XX proteins which are useful for targeting tumour cells associated with  
XX cancer, e.g. adenocarcinomas -

XX Example 5; Page 95; 100pp; English.

XX The present sequence encodes a fusion of an anti-CD20 single chain  
XX antibody and streptavidin. The fusion protein is expressed using  
XX vectors of the invention. The specification describes vector constructs  
XX for expressing streptavidin fusion proteins. The vector comprises a  
XX nucleic acid encoding streptavidin or its functional variant operatively  
XX linked to a promoter, and a cloning site for insertion of a second  
XX nucleic acid sequence encoding a polypeptide to be fused with  
XX streptavidin, interposed between the promoter and the first nucleic  
XX acid sequence. Alternatively, the vector construct comprises a nucleic  
XX acid, operatively linked to a promoter, encoding a polypeptide to be  
XX fused with streptavidin, and a cloning site for insertion of a second  
XX nucleic acid encoding at least 129 amino acids of streptavidin or its  
XX functional variant. The fusion proteins are useful for targeting tumour  
XX cells, particularly tumour cells associated with cancer,  
XX e.g. adenocarcinomas or hematological malignancies. The vector construct

CC is useful for expressing of streptavidin fusion proteins. In particular,  
CC these are useful as tools for medical diagnostics and therapeutic  
CC purposes, e.g. for detecting the presence or absence of, or treating, a  
CC target site within a mammalian host.

XX  
SQ Sequence 765 BP; 170 A; 201 C; 231 G; 163 T; 0 other;

Alignment Scores:  
Pred. No.: 2.33e-09 Length: 765  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.07% Indels: 0  
DB: 22 Gaps: 0

US-08-836-455-4 (1-153) x AAC86591 (1-765)

QY 49 ThrSerTyrAsnMetHisTrpValLysGlnThrProGlyGlnGlyLeuGluTrpIleGly 68  
Db 88 ACCAGTTACAAATATGCACTGGGTAAACGACACCTGGACAGGCCCTGGGATGGATTGGA 147

RESULT 13

AAC86563

ID AAC86563 standard; DNA; 1239 BP.

XX AAC86563;

XX 02-APR-2001 (first entry)

XX DNA encoding a fusion of anti-CD20 single chain antibody/streptavidin.

XX Streptavidin; tumour cell; cancer; adenocarcinoma;  
KW hematological malignancy; ss.

XX Synthetic.

OS Streptomyces avidinii.

OS Homo sapiens.

XX WO200075333-A1.

XX 14-DEC-2000.

XX 05-JUN-2000; 2000WO-US15595.

XX 07-JUN-1999; 99US-0137900.

XX 03-DEC-1999; 99US-0168976.

XX (NEOR-) NEORX CORP.

XX Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;

XX WPI; 2001-091213/10.

XX P-PSDB; AAB30694.

XX New vector constructs for expressing genomic streptavidin fusion  
XX proteins which are useful for targeting tumour cells associated with  
XX cancer, e.g. adenocarcinomas -

XX Example 2; Fig 11A; 100pp; English.

XX The present sequence encodes a fusion of an anti-CD20 single chain  
XX antibody (B9E9) streptavidin. The fusion protein is expressed using  
XX vectors of the invention. The specification describes vector constructs  
XX for expressing streptavidin fusion proteins. The vector comprises a  
XX nucleic acid encoding streptavidin or its functional variant operatively  
XX linked to a promoter, and a cloning site for insertion of a second  
XX nucleic acid sequence encoding a polypeptide to be fused with  
XX streptavidin, interposed between the promoter and the first nucleic  
XX acid sequence. Alternatively, the vector construct comprises a nucleic  
XX acid, operatively linked to a promoter, encoding a polypeptide to be  
XX fused with streptavidin, and a cloning site for insertion of a second  
XX nucleic acid encoding at least 129 amino acids of streptavidin or its  
XX functional variant. The fusion proteins are useful for targeting tumour

CC cells, particularly tumour cells associated with cancer,  
 CC e.g. adenocarcinomas or hematological malignancies. The vector construct  
 CC is useful for expressing of streptavidin fusion proteins. In particular,  
 CC these are useful as tools for medical diagnostics and therapeutic  
 CC purposes, e.g. for detecting the presence or absence of, or treating, a  
 CC target site within a mammalian host.

XX Sequence 1239 BP; 270 A; 392 C; 356 G; 221 T; 0 other;

Alignment Scores:  
 Pred. No.: 3 64e-09 Length: 1239  
 Score: 20.00 Matches: 20  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 13.07% Indels: 0  
 DB: 22 Gaps: 0

US-08-836-455-4 (1-153) x AAC86563 (1-1239)

QY 49 ThrSeryrAsnMethIstPVallySGlnThrProGlyGlnGlyLeuGluTrpIleGly 68  
 |||||  
 Db 469 ACCAGTTACAATATGCTACTGGGTAAAGCAGACACCTGGACAGGGCCTGGGAATGGATTGGA 528

## RESULT 14

AAC86564  
 ID AAC86564 standard; DNA; 1280 BP.

XX AAC86564;

DT 02-APR-2001 (first entry)

XX Anti-CD20 single chain antibody/streptavidin fusion protein cassette.

XX Streptavidin; tumour cell; cancer; adenocarcinoma;  
 KW hematological malignancy; ss.

XX Synthetic.

OS Streptomyces avidinii.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 3..1274

FT /tag= a

FT /product= "anti-CD20 scFv and streptavidin fusion"

XX WO200075333-A1.

PN 14-DEC-2000.

XX 05-JUN-2000; 2000WO-US15595.

XX 07-JUN-1999; 99US-0137900.

PR 03-DEC-1999; 99US-0168976.

XX (NEOR-) NEORX CORP.

XX Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;  
 PI WPI: 2001-091213/10.

XX P-PSDB; AAB30695.

XX New vector constructs for expressing genomic streptavidin fusion  
 PT proteins which are useful for targeting tumour cells associated with  
 PT cancer, e.g. adenocarcinomas -

XX Example 2; Fig 11C; 100pp; English.

XX The present sequence encodes a fusion of an anti-CD20 single chain  
 CC antibody (B9E9) streptavidin. The fusion protein is expressed using  
 CC vectors of the invention. The specification describes vector constructs  
 CC for expressing streptavidin fusion proteins. The vector comprises a  
 CC nucleic acid encoding streptavidin or its functional variant operatively  
 CC linked to a promoter, and a cloning site for insertion of a second

CC nucleic acid sequence encoding a polypeptide to be fused with  
 CC streptavidin, interposed between the promoter and the first nucleic  
 CC acid sequence. Alternatively, the vector construct comprises a nucleic  
 CC acid operatively linked to a promoter, encoding a polypeptide to be  
 CC fused with streptavidin, and a cloning site for insertion of a second  
 CC nucleic acid encoding at least 129 amino acids of streptavidin or its  
 CC functional variant. The fusion proteins are useful for targeting tumour  
 CC cells, particularly tumour cells associated with cancer,  
 CC e.g. adenocarcinomas or hematological malignancies. The vector construct  
 CC is useful for expressing of streptavidin fusion proteins. In particular,  
 CC these are useful as tools for medical diagnostics and therapeutic  
 CC purposes, e.g. for detecting the presence or absence of, or treating, a  
 CC target site within a mammalian host.

XX Sequence 1280 BP; 267 A; 397 C; 388 G; 228 T; 0 other;

## Alignment Scores:

Pred. No.: 3 75e-09 Length: 1280  
 Score: 20.00 Matches: 20  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 13.07% Indels: 0  
 DB: 22 Gaps: 0

US-08-836-455-4 (1-153) x AAC86564 (1-1280)

QY 49 ThrSeryrAsnMethIstPVallySGlnThrProGlyGlnGlyLeuGluTrpIleGly 68  
 |||||  
 Db 96 ACCAGTTACAATATGCTACTGGGTAAAGCAGACACCTGGACAGGGCCTGGGAATGGATTGGA 155

## RESULT 15

AAA15019

ID AAA15019 standard; DNA; 1925 BP.

XX AAA15019;

XX 21-AUG-2000 (first entry)

XX DNA encoding a CD-20 specific chimeric receptor.

XX CD20-specific receptor; CD-20 specific redirected T cell; leukemia;

KW CD20+ malignancy; non-Hodgkin's lymphoma; myeloablative chemotherapy;

XX stem cell rescue; autoimmune disease; lupus; rheumatoid arthritis; ss.

OS Synthetic.

OS Mus sp.

XX Key Location/Qualifiers

FT CDS 27..1928

FT /\*tag= a

XX WO200023573-A2.

PN 27-APR-2000.

XX 20-OCT-1999; 99WO-US24484.

PR 20-OCT-1998; 98US-0105014.

XX (CITY ) CITY OF HOPE.

XX Raubitschek A, Jensen MC, Wu AM;

XX WPI: 2000-339676/29.

DR P-PSDB; AAY84965.

XX Genetically engineered CD20-specific redirected T cells useful for  
 PT treating a CD20+ malignancy, such as non-Hodgkin's lymphoma or CD20+  
 PT acute or chronic leukemia, and autoimmune disease -

XX Example 1; Page 50-53; 58pp; English.

XX The present sequence encodes a synthetic CD20-specific chimeric

CC receptor. The specification describes CD-20 specific redirected T cells  
CC which express and bear on the cell surface membrane a CD20-chimeric  
CC receptor comprising an intracellular signalling domain, a transmembrane  
CC domain and an extracellular domain, the extracellular domain comprising  
CC a CD20-specific receptor. The genetically engineered CD20-specific  
CC redirected T cells are useful for treating a CD20+ malignancy, such  
CC as non-Hodgkin's lymphoma or CD20+ acute or chronic leukemia, in a  
CC human patient having previously undergone myeloablative chemotherapy and  
CC stem cell rescue. The genetically engineered CD20-specific redirected  
CC T cells are also useful for abrogating an untoward B cell function, such  
CC as autoimmune disease (lupus or rheumatoid arthritis) in a patient.  
XX  
SQ Sequence 1925 BP; 471 A; 554 C; 541 G; 359 T; 0 other;

Alignment Scores:  
Pred. No.: 5,48e-09 Length: 1925  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.07% Indels: 0  
DB: 21 Gaps: 0  
US-08-836-455-4 (1-153) x AAA15019 (1-1925)  
Qy 49 ThrSerTyrAsnMetHisTrpValLysGlnThrProGlyGlnGlyLeuGluTrpIleGly 68  
Db 546 ACCAGTTACAATATGCACCTGGTAAAGCAGACACCTGGACAGGCGCTGGATGGATTGGA 605

Search completed: August 31, 2003, 01:16:26  
Job time : 232.013 secs

...

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 31, 2003, 01:02:06 ; Search time 202.802 Seconds  
(without alignments)  
1730.853 Million cell updates/sec

Title: us-08-836-455-4

Perfect score: 153

Sequence: 1 MNCWVFLFLSITTVGHVSQ.....TVSSAKTTPPPVPLVPGSL 153

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1533700 seqs, 1147125425 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3063161

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh  
-O=/cgn2.1/USPTO.spool/US08836455/runat\_29082003\_132954\_23062/app\_query.fasta\_1.654  
-DB=Published\_Applications\_NA\_QFMT-fastcap -SUFFIX=p2noli.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -PCTO=ALIGN-200 -THR\_SCORE=quality -THR\_MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=DO -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US08836455 -CGN\_1.1\_271 -runat\_29082003\_132954\_23062  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOP=6 -FGAPEXT=7 -XGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA:

- 1: /cgn2.6/ptodata/2/pubpna/US07\_PUBCOMB.seq:
- 2: /cgn2.6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:
- 3: /cgn2.6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:
- 4: /cgn2.6/ptodata/2/pubpna/US06\_PUBCOMB.seq:
- 5: /cgn2.6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:
- 6: /cgn2.6/ptodata/2/pubpna/US07\_PUBCOMB.seq:
- 7: /cgn2.6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:
- 8: /cgn2.6/ptodata/2/pubpna/US08\_PUBCOMB.seq:
- 9: /cgn2.6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:
- 10: /cgn2.6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:
- 11: /cgn2.6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:
- 12: /cgn2.6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:
- 13: /cgn2.6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:
- 14: /cgn2.6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:
- 15: /cgn2.6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:
- 16: /cgn2.6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:
- 17: /cgn2.6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	153	100.0	461	9	US-09-861-294-3
					Sequence 3, Appl1

2	153	100.0	461	12	US-10-367-506-3
3	31	20.3	458	10	US-09-990-205-3
4	31	20.3	458	14	US-10-153-401-3
5	20	13.1	765	12	US-10-244-821-44
6	20	13.1	765	12	US-10-244-821-45
7	20	13.1	765	14	US-10-013-173-44
8	20	13.1	765	14	US-10-013-173-45
9	20	13.1	765	14	US-10-150-762-44
10	20	13.1	765	14	US-10-150-762-45
11	20	13.1	1239	14	US-10-244-821-5
12	20	13.1	1239	14	US-10-013-173-5
13	20	13.1	1239	14	US-10-150-762-5
14	20	13.1	1280	12	US/10/244
15	20	13.1	1280	14	US/10/013
16	20	13.1	1280	14	US/10/150
17	19	12.4	736	13	US-10-006-773-12
18	19	12.4	786	12	US-10-244-821-37
19	19	12.4	786	14	US-10-013-173-37
20	19	12.4	786	14	US-10-150-762-37
21	17	11.1	414	13	US-10-146-305-6
22	17	11.1	454	10	US-09-797-941A-1
23	17	11.1	518	9	US-09-881-823-17
24	16	10.5	360	10	US-09-910-059-10
25	16	10.5	408	9	US-09-564-329A-10
26	16	10.5	408	10	US-09-855-153-10
27	16	10.5	408	10	US-09-854-811-10
28	16	10.5	408	10	US-09-934-773-10
29	16	10.5	408	10	US-09-963-620-10
30	16	10.5	408	11	US-09-855-632-10
31	16	10.5	408	12	US-10-224-720-10
32	16	10.5	408	12	US-10-225-779-10
33	16	10.5	408	14	US-10-225-784-10
34	16	10.5	478	13	US-10-040-739-911
35	16	10.5	765	10	US-09-910-059-18
36	16	10.5	1173	12	US-10-075-947A-4
37	15	9.8	339	10	US-09-924-099-12
38	15	9.8	412	10	US-09-924-099-28
39	15	9.8	436	11	US-09-967-719C-7
40	15	9.8	462	11	US-09-967-719C-1
41	15	9.8	690	14	US-09-995-693-4
42	15	9.8	690	14	US-10-232-408-4
43	15	9.8	711	10	US-09-924-099-19
44	15	9.8	714	8	US-08-940-544-4
45	15	9.8	714	12	US-10-075-947A-2

ALIGNMENTS

RESULT 1  
US-09-861-294-3  
; Sequence 3, Application US/09861294  
; Patent No. US20020098190A1  
; GENERAL INFORMATION:  
; APPLICANT: Malaya CHATTERJEE  
; APPLICANT: Kenneth A. FOON  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TUMORS BEARING HMFG AND CEA ANTIGENS  
; FILE REFERENCE: 304142000620  
; CURRENT APPLICATION NUMBER: US/09/861,294  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/049,540  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: 09/096,244  
; PRIOR FILING DATE: 1998-06-11  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 3  
; LENGTH: 461  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(461)

```
; NAME/KEY: sig_peptide
; LOCATION: (1)...(57)
; NAME/KEY: mat_peptide
; LOCATION: (58)...(461)
US-09-861-294-3

Alignment Scores:
Pred. No.: 6,85e-146 Length: 461
Score: 153.00 Matches: 153
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-08-836-455-4 (1-153) x US-09-861-294-3 (1-461)
QY 1 MetGluCysSerTrpValPheLeuLeuSerIleThrThrGlyValHisSerGln 20
Db 1 ATGGAATGCAGCTGGGCTTCTCTCTCTGTCATAACTACAGGTGTCCACTCCAG 60
QY 21 AlaTyrLeuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 61 GCTTATCTACAGCAGTCTGGGCTGAGCTGTGGGCTCAGTGAAGATGCTC 120
QY 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstTrpValLysGlnThrPro 60
Db 121 TGCAAGGCTTCTGGCTACACATTGACCATGACATGACATGCTGGTAAAGCAGACACCT 180
QY 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80
Db 181 GGACAGGCTTGGTACATGATGGAATATTTTCTCTGGAATATTTTCTCTGGAATGCTGATACCAAT 240
QY 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100
Db 241 CAGAAGTTTAAAGGCAAGGCTCATTCACATGCAGACACATCCTCCAGCAGCCATCATG 300
QY 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaAaGlyAsnTrp 120
Db 301 CAGATCAGCAGCTGACATCTGAAGACTCTGGGCTCTATTCTGTGCAAGAGGAACTGG 360
QY 121 GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThr 140
Db 361 GAGGCTCTGGGCTACTGGGCTCAAGAACTCAGTCAACCGCTCTCTCAGCCAAAACG 420

QY 141 ThrProProValTyrProLeuValProGlySerLeu 153
Db 421 ACACCCCAACCGCTCTATCCACTGGTCCCTGGAAGCTTG 459

RESULT 2
US-10-367-506-3
; Sequence 3, Application US/10367506
; Publication No. US20030152575A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth A. FOON
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: TUMORS BEARING HMGF AND CEA ANTIGENS
; FILE REFERENCE: 304142000620
; CURRENT APPLICATION NUMBER: US/10/367,506
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US/09/861,294
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/049,540
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 09/096,244
; PRIOR FILING DATE: 1998-06-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:

; NAME/KEY: CDS
; LOCATION: (1)...(461)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (1)...(57)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (58)...(461)
US-10-367-506-3

Alignment Scores:
Pred. No.: 6,85e-146 Length: 461
Score: 153.00 Matches: 153
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-08-836-455-4 (1-153) x US-10-367-506-3 (1-461)
QY 1 MetGluCysSerTrpValPheLeuLeuSerIleThrThrGlyValHisSerGln 20
Db 1 ATGGAATGCAGCTGGGCTTCTCTCTCTGTCATAACTACAGGTGTCCACTCCAG 60
QY 21 AlaTyrLeuGlnGlnSerGlyAlaGluLeuValArgSerGlyAlaSerValLysMetSer 40
Db 61 GCTTATCTACAGCAGTCTGGGCTGAGCTGTGGGCTCAGTGAAGATGCTC 120
QY 41 CysLysAlaSerGlyTyrThrLeuThrSerTyrAsnMethIstTrpValLysGlnThrPro 60
Db 121 TGCAAGGCTTCTGGCTACACATTGACCATGACATGACATGCTGGTAAAGCAGACACCT 180
QY 61 GlyGlnGlyLeuGluTrpIleGlyAsnIlePheProGlyAsnGlyAspThrTyrTyrAsn 80
Db 181 GGACAGGCTTGGTACATGATGGAATATTTTCTCTGGAATATTTTCTCTGGAATGCTGATACCAAT 240
QY 81 GlnLysPheLysGlyLysAlaSerLeuThrAlaAspThrSerSerSerThrAlaTyrMet 100
Db 241 CAGAAGTTTAAAGGCAAGGCTCATTCACATGCAGACACATCCTCCAGCAGCCATCATG 300
QY 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaAaGlyAsnTrp 120
Db 301 CAGATCAGCAGCTGACATCTGAAGACTCTGGGCTCTATTCTGTGCAAGAGGAACTGG 360
QY 121 GluGlyAlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThr 140
Db 361 GAGGCTCTGGGCTACTGGGCTCAAGAACTCAGTCAACCGCTCTCTCAGCCAAAACG 420

QY 141 ThrProProValTyrProLeuValProGlySerLeu 153
Db 421 ACACCCCAACCGCTCTATCCACTGGTCCCTGGAAGCTTG 459

RESULT 3
US-09-990-205-3
; Sequence 3, Application US/09990205
; Patent No. US20020150572A1
; GENERAL INFORMATION:
; APPLICANT: FOON, Kenneth A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF PSORIASIS
; FILE REFERENCE: 304142000501
; CURRENT APPLICATION NUMBER: US/09/990,205
; CURRENT FILING DATE: 2001-11-20
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: U.S. 60/065,774
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Mus Musculus
```



```
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(456)
; NAME/KEY: mat_peptide
; LOCATION: (58)..(456)
US-09-990-205-3

Alignment Scores:
Pred. No.: 2,46e-22 Length: 458
Score: 31.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.26% Indels: 0
DB: 10 Gaps: 0

US-08-836-455-4 (1-153) x US-09-990-205-3 (1-458)
QY 123 AlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrThrPro 142
Db 364 GCTCTGGACTACTGGGGTCAAGGACCTCAGTCACCGTCTCTCCAGCAAAAGACACCC 423
QY 143 ProProValTyrProLeuValProGlySerLeu 153
Db 424 CCACCGCTCTATCCATGGTCCCTCGAAGCTTG 456

RESULT 4
US-10-153-401-3
; Sequence 3, Application US/10153401
; Publication No. US20030114398A1
; GENERAL INFORMATION:
; APPLICANT: Chatterjee, Malay A
; Foon, Kenneth A.
; TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE
; TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/153,401
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/293,533
; FILING DATE: 1999-04-15
; APPLICATION NUMBER: US 08/372,676
; FILING DATE: 1995-01-17
; APPLICATION NUMBER: US 08/591,196
; FILING DATE: 1996-01-16
; ATTORNEY/AGENT INFORMATION:
; NAME: Catherine M. Polizzi
; REGISTRATION NUMBER: 40,130
; REFERENCE/DOCKET NUMBER: 304142000202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 458 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
;
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..456
; NAME/KEY: mat_peptide
; LOCATION: 58
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-153-401-3

Alignment Scores:
Pred. No.: 2,46e-22 Length: 458
Score: 31.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.26% Indels: 0
DB: 14 Gaps: 0

US-08-836-455-4 (1-153) x US-10-153-401-3 (1-458)
QY 123 AlaLeuAspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrThrPro 142
Db 364 GCTCTGGACTACTGGGGTCAAGGACCTCAGTCACCGTCTCTCCAGCAAAAGACACCC 423
QY 143 ProProValTyrProLeuValProGlySerLeu 153
Db 424 CCACCGCTCTATCCATGGTCCCTCGAAGCTTG 456

RESULT 5
US-10-244-821-44
; Sequence 44, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yunkang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic Streptavidin
; OTHER INFORMATION: fusion construct
US-10-244-821-44

Alignment Scores:
Pred. No.: 5.37e-11 Length: 765
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.07% Indels: 0
DB: 12 Gaps: 0

US-08-836-455-4 (1-153) x US-10-244-821-44 (1-765)
QY 49 ThrSerTyrAsnMethHisTrpValLysGlnThrProGlyGlnGlyLeuGluTrpIleGly 68
Db 88 ACCAGTTACAATATGCACTGGTAAAGCAGACACCTGGAGCGCCTCGAATGGATTGA 147

RESULT 6
US-10-244-821-45
; Sequence 45, Application US/10244821
```

```

; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstynne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C2B8 single chain antibody-genomic Streptavidin
; OTHER INFORMATION: fusion construct
US-10-244-821-45

Alignment Scores:
Pred. No.: 5.37e-11 Length: 765
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.07% Indels: 0
DB: 12 Gaps: 0

US-08-836-455-4 (1-153) x US-10-244-821-45 (1-765)

Qy 49 ThrSerTyrAsnMetHisTrpValLysGlnThrProGlyGlnGlyLeuGluTrpIleGly 68
Db 88 ACCAGTTACAATATGCACTGGTAAAGCAGACACCTGGAGGCGCTGGATGATGGA 147

RESULT 7
US-10-013-173-44
; Sequence 44, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic Streptavidin
; OTHER INFORMATION: fusion construct
US-10-013-173-44

Alignment Scores:
Pred. No.: 5.37e-11 Length: 765
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.07% Indels: 0
DB: 14 Gaps: 0

US-08-836-455-4 (1-153) x US-10-013-173-44 (1-765)

Qy 49 ThrSerTyrAsnMetHisTrpValLysGlnThrProGlyGlnGlyLeuGluTrpIleGly 68
Db 88 ACCAGTTACAATATGCACTGGTAAAGCAGACACCTGGAGGCGCTGGATGATGGA 147

RESULT 8
US-10-013-173-45
; Sequence 45, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C2B8 single chain antibody-genomic Streptavidin
; OTHER INFORMATION: fusion construct
US-10-013-173-45

Alignment Scores:
Pred. No.: 5.37e-11 Length: 765
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.07% Indels: 0
DB: 14 Gaps: 0

US-08-836-455-4 (1-153) x US-10-013-173-45 (1-765)

Qy 49 ThrSerTyrAsnMetHisTrpValLysGlnThrProGlyGlnGlyLeuGluTrpIleGly 68
Db 88 ACCAGTTACAATATGCACTGGTAAAGCAGACACCTGGAGGCGCTGGATGATGGA 147

RESULT 9
US-10-150-762-44
; Sequence 44, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; APPLICANT: Dearstynne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150,762
; CURRENT FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic Streptavidin
; OTHER INFORMATION: fusion construct
US-10-013-173-44

Alignment Scores:
Pred. No.: 5.37e-11 Length: 765
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.07% Indels: 0
DB: 14 Gaps: 0
```

```

US-08-836-455-4 (1-153) x US-10-013-173-44 (1-765)

Qy 49 ThrSerTyrAsnMetHisTrpValLysGlnThrProGlyGlnGlyLeuGluTrpIleGly 68
Db 88 ACCAGTTACAATATGCACTGGTAAAGCAGACACCTGGAGGCGCTGGATGATGGA 147

RESULT 8
US-10-013-173-45
; Sequence 45, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C2B8 single chain antibody-genomic Streptavidin
; OTHER INFORMATION: fusion construct
US-10-013-173-45

Alignment Scores:
Pred. No.: 5.37e-11 Length: 765
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.07% Indels: 0
DB: 14 Gaps: 0

US-08-836-455-4 (1-153) x US-10-013-173-45 (1-765)

Qy 49 ThrSerTyrAsnMetHisTrpValLysGlnThrProGlyGlnGlyLeuGluTrpIleGly 68
Db 88 ACCAGTTACAATATGCACTGGTAAAGCAGACACCTGGAGGCGCTGGATGATGGA 147

RESULT 9
US-10-150-762-44
; Sequence 44, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; APPLICANT: Dearstynne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150,762
; CURRENT FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic Streptavidin
; OTHER INFORMATION: fusion construct
US-10-013-173-44

Alignment Scores:
Pred. No.: 5.37e-11 Length: 765
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.07% Indels: 0
DB: 14 Gaps: 0
```

```
; OTHER INFORMATION: fusion construct
US-10-150-762-44
Alignment Scores:
Pred. No.: 5,37e-11 Length: 765
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.07% Indels: 0
DB: 14 Gaps: 0

US-08-836-455-4 (1-153) x US-10-150-762-44 (1-765)
QY 49 ThrSerTyraSmMethIstRpVallysGlnThrProGlyGlnGlyLeuGluTrpIlegly 68
Db 88 ACCAGTTACAATATGCACTGGGTAAAGACAGACACCTGGAGCGGCTGGGAATGGATTGGA 147

RESULT 10
US-10-150-762-45
; Sequence 45, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150,762
; CURRENT FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C2B8 single chain antibody-genomic Streptavidin
US-10-150-762-45
Alignment Scores:
Pred. No.: 5,37e-11 Length: 765
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.07% Indels: 0
DB: 14 Gaps: 0

US-08-836-455-4 (1-153) x US-10-150-762-45 (1-765)
QY 49 ThrSerTyraSmMethIstRpVallysGlnThrProGlyGlnGlyLeuGluTrpIlegly 68
Db 88 ACCAGTTACAATATGCACTGGGTAAAGACAGACACCTGGAGCGGCTGGGAATGGATTGGA 147

RESULT 11
US-10-244-821-5
; Sequence 5, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
US-10-244-821-5
Alignment Scores:
Pred. No.: 8,25e-11 Length: 1239
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.07% Indels: 0
DB: 14 Gaps: 0

US-08-836-455-4 (1-153) x US-10-244-821-5 (1-1239)
QY 49 ThrSerTyraSmMethIstRpVallysGlnThrProGlyGlnGlyLeuGluTrpIlegly 68
Db 469 ACCAGTTACAATATGCACTGGGTAAAGACAGACACCTGGAGCGGCTGGGAATGGATTGGA 528

RESULT 12
US-10-013-173-5
; Sequence 5, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic streptavidin fusion
US-10-013-173-5
Alignment Scores:
Pred. No.: 8,25e-11 Length: 1239
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.07% Indels: 0
DB: 14 Gaps: 0

US-08-836-455-4 (1-153) x US-10-013-173-5 (1-1239)
QY 49 ThrSerTyraSmMethIstRpVallysGlnThrProGlyGlnGlyLeuGluTrpIlegly 68
Db 469 ACCAGTTACAATATGCACTGGGTAAAGACAGACACCTGGAGCGGCTGGGAATGGATTGGA 528

RESULT 13
US-10-150-762-5
; Sequence 5, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
```

```

; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, Jonh M.
; APPLICANT: Dearstynne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150,762
; CURRENT FILING DATE: 2002-05-17
; NUMBER OF SEQ. ID NOS: 90
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic streptavidin fusion
US-10-150-762-5

Alignment Scores:
Pred. No.:      8.25e-11      Length:      1239
Score:          20.00         Matches:      20
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:    13.07%       Indels:       0
DB:             14           Gaps:        0

US-08-836-455-4 (1-153) x US-10-150-762-5 (1-1239)

QY 49 ThrSerTyrrAsnMethIstTpValLysGlnThrProGlyGlnGlyLeuGluTrpIleGly 68
Db 469 ACCAGTTACAAATGTCACCTGGGTAAACACACACACTGGACAGGCCCTGGAAATGGATTGGA 528

RESULT 14
US/10/244.
; Sequence 7, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstynne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1280
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic streptavidin fusion construct
US/10/244,821-7

Alignment Scores:
Pred. No.:      8.5e-11      Length:      1280
Score:          20.00         Matches:      20
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:    13.07%       Indels:       0
DB:             12           Gaps:        0

US-08-836-455-4 (1-153) x US/10/244 (1-1280)

QY 49 ThrSerTyrrAsnMethIstTpValLysGlnThrProGlyGlnGlyLeuGluTrpIleGly 68

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 30, 2003, 22:20:21 ; Search time 1841.13 Seconds  
(without alignments)  
2019.725 Million cell updates/sec

Title: US-08-836-455-4  
Perfect score: 153  
Sequence: 1 MECSWVFLLSITGVHSQ.....TVSSAKTTPPPVPLVPGSL 153

Scoring table:  
OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 1

Total number of hits satisfying chosen parameters: 45562604

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q/cgu2.1/USPTO\_spool/US08836455/runat\_29082003\_132953\_23055/app\_query.fasta\_1.654  
-DB=EST -QFMT=fastap -SUFFIX=p2noli.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.coi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US08836455@cgn\_1.1\_3596/runat\_29082003\_132953\_23055 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :  
EST::\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estcov:\*  
6: em\_estcpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	24	15.7	745	12	BG871607	BG871607 602790090
2	23	15.0	958	13	BUS23796	BUS23796 AGENCOURT
3	21	13.7	749	12	BI096997	BI096997 602884553
4	19	12.4	676	10	BE369087	BE369087 601221765
5	18	11.8	903	12	BG963063	BG963063 602828076
6	17	11.1	724	12	BG962137	BG962137 602826902
7	17	11.1	802	10	BF137216	BF137216 601784466
8	17	11.1	861	12	BI455668	BI455668 603173862
9	17	11.1	862	10	BF143948	BF143948 601786493
10	17	11.1	883	12	BI151077	BI151077 602917012
11	17	11.1	914	10	BF163883	BF163883 601772396
12	17	11.1	1012	10	BF142302	BF142302 601791844
13	16	10.5	393	10	BG095410	BG095410 mac28d12.
14	16	10.5	439	10	BE136816	BE136816 ug56b01.y
15	16	10.5	529	14	CA577954	CA577954 K0716F08-
16	16	10.5	639	10	BE371136	BE371136 601218628
17	16	10.5	829	10	BF144014	BF144014 601791270
18	16	10.5	861	12	BG962366	BG962366 602827176
19	16	10.5	979	10	BF578874	BF578874 602095434
c 20	15	9.8	124	13	BQ358790	BQ358790 MK1-HT106
c 21	15	9.8	136	12	BI043088	BI043088 QV1-OT017
c 22	15	9.8	176	10	BF822838	BF822838 MK1-SN007
c 23	15	9.8	212	13	BQ364985	BQ364985 MK4-SN007
c 24	15	9.8	379	10	AW908776	AW908776 uf57d02.y
c 25	15	9.8	389	9	AI555192	AI555192 UI-R-C2p-
26	15	9.8	417	14	CA577519	CA577519 K0710F01-
27	15	9.8	462	14	CA578969	CA578969 K0731A05-
28	15	9.8	481	14	CA576938	CA576938 K0702C06-
29	15	9.8	509	14	CA578347	CA578347 K0722D09-
30	15	9.8	517	10	BF545108	BF545108 UI-R-C2p-
31	15	9.8	598	12	BI149725	BI149725 602848572
32	15	9.8	607	9	AV259014	AV259014 AV259014
33	15	9.8	616	10	BF136295	BF136295 601781413
34	15	9.8	637	14	BY724721	BY724721 BY724721
35	15	9.8	644	10	BE281961	BE281961 601099518
36	15	9.8	689	11	AK007163	AK007163 Mus muscu
37	15	9.8	718	10	BF136279	BF136279 601780988
38	15	9.8	739	10	BE284158	BE284158 601099428
39	15	9.8	748	10	BF136397	BF136397 601783927
40	15	9.8	794	12	BI150371	BI150371 602915205
41	15	9.8	801	13	BU946353	BU946353 AGENCOURT
42	15	9.8	853	12	BI249876	BI249876 602995978
43	15	9.8	891	10	BF138460	BF138460 601782916
44	15	9.8	892	10	BG518664	BG518664 602578528
45	15	9.8	896	12	BI149320	BI149320 602848859

ALIGNMENTS

RESULT 1  
BG871607  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

BG871607 602790090F1 NCI\_CGAP\_SG2 Mus musculus cdna clone IMAGE:4921295 5',  
745 bp mRNA linear EST 29-MAY-2001  
mRNA sequence.  
BG871607.1 GI:14222147  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 745)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1AM10839 row: c column: 24  
 High quality sequence stop: 582.  
 Location/Qualifiers  
 1. .745  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clones="IMAGE:4921295"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_SG2"  
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 179 a 222 c 181 g 163 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.37e-12 Length: 745  
 Score: 24.00 Matches: 24  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 15.69% Indels: 0  
 DB: 12 Gaps: 0

US-08-836-455-4 (1-153) x BG871607 (1-745)

Qy 94 SerSerThrAlaTyrMetGlnIleSerSerLeuThrSerGluAspSerAlaValTyr 113  
 |||||||  
 Db 315 TCCTCCAGCAGCGCTACATGCAGATCAGCAGCTGACCTCTGAGGACTCTGCGGTCTAT 374  
 Qy 114 PheCysalaArg 117  
 |||||||  
 Db 375 TTCTGTGCAAGA 386

RESULT 2  
 BU523796  
 LOCUS BU523796 958 bp mRNA linear EST 13-SEP-2002  
 DEFINITION AGENCOURT\_10123257 NCI\_CGAP\_Co24 Mus musculus cDNA clone  
 IMAGE:6530504 5', mRNA sequence.  
 ACCESSION BU523796  
 VERSION BU523796.1 GI:22834235  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 958)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1AM10839 row: c column: 24  
 High quality sequence stop: 582.  
 Location/Qualifiers  
 1. .745  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clones="IMAGE:4921295"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_SG2"  
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 179 a 222 c 181 g 163 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.37e-12 Length: 745  
 Score: 24.00 Matches: 24  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 15.69% Indels: 0  
 DB: 12 Gaps: 0

US-08-836-455-4 (1-153) x BG871607 (1-745)

Qy 94 SerSerThrAlaTyrMetGlnIleSerSerLeuThrSerGluAspSerAlaValTyr 113  
 |||||||  
 Db 315 TCCTCCAGCAGCGCTACATGCAGATCAGCAGCTGACCTCTGAGGACTCTGCGGTCTAT 374  
 Qy 114 PheCysalaArg 117  
 |||||||  
 Db 375 TTCTGTGCAAGA 386

RESULT 2  
 BU523796  
 LOCUS BU523796 958 bp mRNA linear EST 13-SEP-2002  
 DEFINITION AGENCOURT\_10123257 NCI\_CGAP\_Co24 Mus musculus cDNA clone  
 IMAGE:6530504 5', mRNA sequence.  
 ACCESSION BU523796  
 VERSION BU523796.1 GI:22834235  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 958)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1AM10839 row: c column: 24  
 High quality sequence stop: 582.  
 Location/Qualifiers  
 1. .745  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clones="IMAGE:4921295"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_SG2"  
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 179 a 222 c 181 g 163 t  
 ORIGIN

Plate: L1AM14132 row: f column: 08  
 High quality sequence stop: 608.  
 Location/Qualifiers  
 1. .958  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clones="IMAGE:6530504"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Co24"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 233 a 285 c 231 g 208 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.57e-11 Length: 958  
 Score: 23.00 Matches: 23  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 15.03% Indels: 0  
 DB: 13 Gaps: 0

US-08-836-455-4 (1-153) x BU523796 (1-958)

Qy 94 SerSerThrAlaTyrMetGlnIleSerSerLeuThrSerGluAspSerAlaValTyr 113  
 |||||||  
 Db 318 TCCTCCAGCAGCGCTACATGCAGATCAGCAGCTGACCTCTGAGGACTCTGCGGTCTAT 377  
 Qy 114 PheCysala 116  
 |||||||  
 Db 378 TTCTGTGCA 386

RESULT 3  
 BU099697  
 LOCUS BU099697 749 bp mRNA linear EST 26-JUN-2001  
 DEFINITION 60288453F1 NCI\_CGAP\_Kid14 Mus musculus cDNA clone IMAGE:5040084  
 5', mRNA sequence.  
 ACCESSION BU099697  
 VERSION BU099697.1 GI:14550590  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 749)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1AM1110 row: i column: 13  
 High quality sequence stop: 741.  
 Location/Qualifiers  
 1. .749  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clones="IMAGE:5040084"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Kid14"  
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site\_1: NotI;

Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library. I"

BASE COUNT 183 a 214 c 185 g 167 t  
ORIGIN

Alignment Scores:  
Pred. No.: 9.92e-10 Length: 749  
Score: 21.00 Matches: 21  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.73% Indels: 0  
DB: 12 Gaps: 0

US-08-836-455-4 (1-153) x BT099697 (1-749)

Qy 97 ThrAlaTyrMetGlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAla 116  
|||||  
Db 118 ACAGCTACATGACATGACAGCTGCTGAGACACTCTGCGGTCTATTCTGTGGC 177  
Qy 117 Arg 117  
Db 178 AGA 180

RESULT 4  
BE369087 601221765F1 NCI\_CGAP\_Lu29 Mus musculus cdna clone IMAGE:3590320 5',  
LOCUS mRNA sequence.  
DEFINITION

ACCESSION BE369087

VERSION BE369087.1 GI:9314450

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 676)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM8758 row: b column: 17

High quality sequence stop: 580.

Location/Qualifiers

1..676

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clones="IMAGE:3590320"

/tissue\_type="spontaneous tumor, metastatic to mammary."

Stem cell origin."

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP\_Lu29"

/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: SalI;

Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

BASE COUNT 161 a 188 c 173 g 154 t

ORIGIN

Alignment Scores:

Pred. No.: 7.2e-08 Length: 676

Score: 19.00 Matches: 19

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.42% Indels: 0  
DB: 10 Gaps: 0

US-08-836-455-4 (1-153) x BE369087 (1-676)

Qy 125 AspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrThrProPro 143  
|||||  
Db 392 GACTACTGGGTCAAGGAAGCTCAGTCACCGTCTCCCTCAGCAAAAGACACCCCA 448

RESULT 5

BE963063

LOCUS

DEFINITION 602828076F1 NCI\_CGAP\_Co24 Mus musculus cdna clone IMAGE:4982827 5',  
mRNA sequence.

ACCESSION BE963063

VERSION BE963063.1 GI:14350700

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 903)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10986 row: g column: 20

High quality sequence stop: 875.

Location/Qualifiers

1..903

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clones="IMAGE:4982827"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NCI\_CGAP\_Co24"

/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;

Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.6 kb. Constructed by Life

Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 208 a 266 c 229 g 200 t

ORIGIN

Alignment Scores:

Pred. No.: 8.55e-07 Length: 903

Score: 18.00 Matches: 18

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 11.76% Indels: 0

DB: 12 Gaps: 0

US-08-836-455-4 (1-153) x BE963063 (1-903)

Qy 101 GlnIleSerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGly 118  
|||||

Db 50 CAGATCAGACGCTGACCTCTGAGACTCTGCGGTCTATTCTGTGCAAGAGGC 103

RESULT 6

BE962137

LOCUS

DEFINITION 602826902F1 NCI\_CGAP\_Co24 Mus musculus cdna clone IMAGE:4981585 5',  
mRNA sequence.

ACCESSION BE962137

```

VERSION  BG962137.1  GI:14349774
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus

REFERENCE 1 (bases 1 to 724)
AUTHORS  NIH-MGC http://mgc.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
          Tissue Procurement: Jeffrey E. Green, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Cloned through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM10983 row: d column: 02
          High quality sequence stop: 713.

FEATURES  source
          Location/Qualifiers
            1..724
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="FVB/N"
              /db_xref="taxon:10090"
              /clone="IMAGE:4981585"
              /lab_host="DH10B (T1 phage-resistant)"
              /clone_lib="NCI-CCGAP_Co24"
              /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
              Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
              Average insert size 1.6 kb. Constructed by Life
              Technologies. Note: this is a NCI-CCGAP Library."
BASE COUNT  183 a 199 c 180 g 162 t
ORIGIN
    125 AspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrThr 141
    |||||
    435 GACTACTGGGGTCAGGACCTCAGTCACCGCTCTCTCCAGCAACAAACA 485

Alignment Scores:
Pred. NO.: 6.17e-06 Length: 724
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.11% Indels: 0
DB: 12 Gaps: 0

US-08-836-455-4 (1-153) x BG962137 (1-724)

QY 94 SerSerThrAlaTyrMetGlnIleSerSerLeuThrThrSerGluAspSer 110
|||
DB 314 TCCCTCCAGCAGCGCTACATGCAGATCAGCAGCCTGACCTCTGAGGACTCT 364

RESULT 7
BF137216
LOCUS  BF137216 802 bp mRNA linear EST 24-OCT-2000
DEFINITION 601784466P1 NCI-CCGAP_Lu30 Mus musculus cDNA clone IMAGE:4012360 5',
mRNA sequence.
ACCESSION  BF137216
VERSION  BF137216.1 GI:10976256
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          1 (bases 1 to 802)
          NIH-MGC http://mgc.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Gilbert Smith, Ph.D.
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Cloned through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM10983 row: d column: 02
          High quality sequence stop: 713.

FEATURES  source
          Location/Qualifiers
            1..802
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6J"
              /db_xref="taxon:10090"
              /clone="IMAGE:5253279"
              /tissue_type="tumor, metastatic to mammary"
              /lab_host="DH10B"
              /clone_lib="NCI-CCGAP_Lu30"
              /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
              Site_2: SalI; transgenic model WNT-1, expression driven by
              MMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
              dt. Library constructed by Life Technologies.
              Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT  191 a 217 c 211 g 183 t
ORIGIN
    125 AspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrThr 141
    |||||
    435 GACTACTGGGGTCAGGACCTCAGTCACCGCTCTCTCCAGCAACAAACA 485

Cloned through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9253 row: c column: 17
High quality sequence stop: 697.

FEATURES  source
          Location/Qualifiers
            1..802
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6J"
              /db_xref="taxon:10090"
              /clone="IMAGE:4012360"
              /tissue_type="tumor, metastatic to mammary"
              /lab_host="DH10B"
              /clone_lib="NCI-CCGAP_Lu30"
              /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
              Site_2: SalI; transgenic model WNT-1, expression driven by
              MMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
              dt. Library constructed by Life Technologies.
              Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT  191 a 217 c 211 g 183 t
ORIGIN
    125 AspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrThr 141
    |||||
    435 GACTACTGGGGTCAGGACCTCAGTCACCGCTCTCTCCAGCAACAAACA 485

Cloned through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11639 row: l column: 16
High quality sequence stop: 780.

FEATURES  source
          Location/Qualifiers
            1..861
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6J"
              /db_xref="taxon:10090"
              /clone="IMAGE:5253279"
              /tissue_type="tumor, gross tissue"
              /dev_stage="7 months"

```

```

Cloned through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9253 row: c column: 17
High quality sequence stop: 697.

FEATURES  source
          Location/Qualifiers
            1..802
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6J"
              /db_xref="taxon:10090"
              /clone="IMAGE:4012360"
              /tissue_type="tumor, metastatic to mammary"
              /lab_host="DH10B"
              /clone_lib="NCI-CCGAP_Lu30"
              /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
              Site_2: SalI; transgenic model WNT-1, expression driven by
              MMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
              dt. Library constructed by Life Technologies.
              Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT  191 a 217 c 211 g 183 t
ORIGIN
    125 AspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaLysThrThr 141
    |||||
    435 GACTACTGGGGTCAGGACCTCAGTCACCGCTCTCTCCAGCAACAAACA 485

Cloned through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11639 row: l column: 16
High quality sequence stop: 780.

FEATURES  source
          Location/Qualifiers
            1..861
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6J"
              /db_xref="taxon:10090"
              /clone="IMAGE:5253279"
              /tissue_type="tumor, gross tissue"
              /dev_stage="7 months"

```



```

/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam5"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT 205 a 239 c 211 g 206 t
ORIGIN

Alignment Scores:
Pred. No.: 7.31e-06 Length: 861
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.11% Indels: 0
DB: 12 Gaps: 0

US-08-836-455-4 (1-153) x BI455668 (1-861)

QY 125 AspyTrpGlyGlnGlyThrSerValThrValSerSerAlalysThrThr 141
|||||
Db 443 GACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCTCCACCAACAACA 493

RESULT 9
LOCUS BF143948 862 bp mRNA linear EST 24-OCT-2000
DEFINITION 601786493F1 NCI_CGAP_Lu30 Mus musculus cdna clone IMAGE:4014430 5',
mRNA sequence.
ACCESSION BF143948 GI:10982988
VERSION BF143948.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 862)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM9258 row: i column: 23
High quality sequence stop: 671.
FEATURES
Location/Qualifiers
1..862
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4014430"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; transgenic model WNT-1, expression driven by
MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dt. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 213 a 236 c 236 g 176 t
ORIGIN

Alignment Scores:
Pred. No.: 7.32e-06 Length: 862
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0

```

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.11% Indels: 0
DB: 10 Gaps: 0

US-08-836-455-4 (1-153) x BF143948 (1-862)

QY 125 AspyTrpGlyGlnGlyThrSerValThrValSerSerAlalysThrThr 141
|||||
Db 412 GACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCTCCACCAACAACA 462

RESULT 10
LOCUS BF151077 883 bp mRNA linear EST 05-JUL-2001
DEFINITION 602917012F1 NCI_CGAP_Lu29 Mus musculus cdna clone IMAGE:5067314 5',
mRNA sequence.
ACCESSION BF151077
VERSION BF151077.1 GI:14611078
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 883)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM11181 row: h column: 03
High quality sequence stop: 719.
FEATURES
Location/Qualifiers
1..883
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5067314"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 211 a 268 c 217 g 187 t
ORIGIN

Alignment Scores:
Pred. No.: 7.49e-06 Length: 883
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.11% Indels: 0
DB: 12 Gaps: 0

US-08-836-455-4 (1-153) x BF151077 (1-883)

QY 125 AspyTrpGlyGlnGlyThrSerValThrValSerSerAlalysThrThr 141
|||||
Db 257 GACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCTCCACCAACAACA 307

RESULT 11
LOCUS BF163883 914 bp mRNA linear EST 30-OCT-2000
DEFINITION 601772396F1 NCI_CGAP_Lu29 Mus musculus cdna clone IMAGE:3991558 5',

```

```

ACCESSION      Bf163883
VERSION        Bf163883.1 GI:11044161
KEYWORDS       EST.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
REFERENCE      1 (bases 1 to 914)
AUTHORS        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL        NIH-MGC http://mgc.nci.nih.gov/.
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-f@mail.nih.gov
               Tissue Procurement: Gilbert Smith, Ph.D.
               cDNA Library Preparation: Life Technologies, Inc.
               DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L1A9204 row: p column: 23
               High quality sequence stop: 679.
FEATURES       Location/Qualifiers
               source          1..914
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="Czech II"
               /db_xref="taxon:10090"
               /clone="IMAGE:3991558"
               /tissue_type="spontaneous tumor, metastatic to mammary.
               Stem cell origin."
               /lab_host="DH10B"
               /clone_lib="NCI_CGAP_Lu29"
               /note="Organ: lung; Vector: pCMV-SPORT6; Site.1: SalI;
               Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.
               Library constructed by Life Technologies. Investigator
               providing samples: Gilbert Smith, NIH"
BASE COUNT    220 a 228 c 265 g 201 t
ORIGIN
Alignment Scores:
Pred. No.:      7,75e-06          Length:      914
Score:          17.00             Matches:      17
Percent Similarity: 100.00%       Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match:    11.11%           Indels:      0
DB:             10               Gaps:        0

US-08-836-455-4 (1-153) x Bf163883 (1-914)

Qy 125 AspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaIysThrThr 141
|||||
Db 415 GACTACTGGGGTCAAGGACGTGAGTCCGCTCTCTCCAGCAACAAACA 465

RESULT 12
Bf142302
LOCUS          Bf142302.1 GI:10981252
DEFINITION    601791844F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4022459 5',
               mRNA sequence.
ACCESSION     Bf142302
VERSION       Bf142302.1 GI:10981252
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE      1 (bases 1 to 1012)
AUTHORS        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL        NIH-MGC http://mgc.nci.nih.gov/.
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-f@mail.nih.gov
               Tissue Procurement: Gilbert Smith, Ph.D.
               cDNA Library Preparation: Life Technologies, Inc.
               DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L1A9204 row: p column: 23
               High quality sequence stop: 679.
FEATURES       Location/Qualifiers
               source          1..914
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="Czech II"
               /db_xref="taxon:10090"
               /clone="IMAGE:3991558"
               /tissue_type="spontaneous tumor, metastatic to mammary.
               Stem cell origin."
               /lab_host="DH10B"
               /clone_lib="NCI_CGAP_Lu29"
               /note="Organ: lung; Vector: pCMV-SPORT6; Site.1: SalI;
               Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.
               Library constructed by Life Technologies. Investigator
               providing samples: Gilbert Smith, NIH"
BASE COUNT    220 a 228 c 265 g 201 t
ORIGIN
Alignment Scores:
Pred. No.:      7,75e-06          Length:      914
Score:          17.00             Matches:      17
Percent Similarity: 100.00%       Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match:    11.11%           Indels:      0
DB:             10               Gaps:        0

US-08-836-455-4 (1-153) x Bf163883 (1-914)

Qy 125 AspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaIysThrThr 141
|||||
Db 415 GACTACTGGGGTCAAGGACGTGAGTCCGCTCTCTCCAGCAACAAACA 465

RESULT 13
Bf095410
LOCUS          Bf095410.1 GI:12577973
DEFINITION    mac28d12.y1 Soares mouse 3Nbms Mus musculus cDNA clone
               IMAGE:4000967 5' similar to SW:DTCM_MOUSE P01882 IG DELTA CHAIN C
               REGION MEMBRANE-BOUND FORM. ;, mRNA sequence.
ACCESSION     Bf095410
VERSION       Bf095410.1 GI:12577973
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE          1 (bases 1 to 393)
JOURNAL        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
               Unpublished
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-f@mail.nih.gov
               This clone is available royalty-free through LLNL ; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               MGI:1498695
               Seq primer: -40RP from Gibco
               High quality sequence stop: 334.
FEATURES       Location/Qualifiers
               source          1..393
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="C57BL/6J"
               /db_xref="taxon:10090"
               /clone="IMAGE:4000967"

```

```

Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1A9279 row: h column: 12
High quality sequence stop: 711.
FEATURES       Location/Qualifiers
               source          1..1012
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="Czech II"
               /db_xref="taxon:10090"
               /clone="IMAGE:4022459"
               /tissue_type="tumor, metastatic to mammary"
               /lab_host="DH10B"
               /clone_lib="NCI_CGAP_Lu30"
               /note="Organ: lung; Vector: pCMV-SPORT6; Site.1: NotI;
               Site.2: SalI; transgenic model WNT-1, expression driven by
               MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
               dt. Library constructed by Life Technologies.
               Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT    273 a 271 c 267 g 201 t
ORIGIN
Alignment Scores:
Pred. No.:      8,56e-06          Length:      1012
Score:          17.00             Matches:      17
Percent Similarity: 100.00%       Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match:    11.11%           Indels:      0
DB:             10               Gaps:        0

US-08-836-455-4 (1-153) x Bf142302 (1-1012)

Qy 125 AspTyrTrpGlyGlnGlyThrSerValThrValSerSerAlaIysThrThr 141
|||||
Db 413 GACTACTGGGGTCAAGGACGTGAGTCCGCTCTCTCCAGCAACAAACA 463

RESULT 13
Bf095410
LOCUS          Bf095410.1 GI:12577973
DEFINITION    mac28d12.y1 Soares mouse 3Nbms Mus musculus cDNA clone
               IMAGE:4000967 5' similar to SW:DTCM_MOUSE P01882 IG DELTA CHAIN C
               REGION MEMBRANE-BOUND FORM. ;, mRNA sequence.
ACCESSION     Bf095410
VERSION       Bf095410.1 GI:12577973
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE          1 (bases 1 to 393)
JOURNAL        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
               Unpublished
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-f@mail.nih.gov
               This clone is available royalty-free through LLNL ; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               MGI:1498695
               Seq primer: -40RP from Gibco
               High quality sequence stop: 334.
FEATURES       Location/Qualifiers
               source          1..393
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="C57BL/6J"
               /db_xref="taxon:10090"
               /clone="IMAGE:4000967"

```

```

/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mouse 3NDMS"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TCTTACCAATCTGAAGTGGAGCGCGCGCTGTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(AATTGGGATCTTG), digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Source irradiated bowel harvested 72 hours after
irradiation (1400 Gys). Library constructed by Bob
Barstead."
BASE COUNT 108 a 95 c 93 g 97 t
ORIGIN

Alignment Scores:
Pred. No.: 3,04e-05 Length: 393
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.46% Indels: 0
DB: 10 Gaps: 0

US-08-836-455-4 (1-153) x BG095410 (1-393)

QY 103 SerSerLeuthrSerGluAspSerAlaValTyrPheCysAlaArgGly 118
|||||
Db 40 AGCAGCCTGACATCTGAGGACTCTGGGCTCTATTCTGTGCCAGGGGT 87

RESULT 14
BE136816
LOCUS
DEFINITION
ug55b01.y1 Barstead bowel MPLRB9 Mus musculus cDNA clone
IMAGE:1546345 5' similar to gb:M12376 Mouse immunoglobulin H-chain
V-region pseudogene mRNA, complete (MOUSE);, mRNA sequence.
BE136816
BE136816.1 GI:8599316
EST.
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 439)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter
, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LILML; contact the
IMAGE Consortium (info@image.lilml.gov) for further information.
MGI:951693
Seq primer: -40RP from Gibco.
FEATURES
source
1. .439
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1546345"
/tissue_type="bowel"
/dev_stage="8 weeks"
/lab_host="DH10B"

/clone_lib="Barstead bowel MPLRB9"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TCTTACCAATCTGAAGTGGAGCGCGCGCTGTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(AATTGGGATCTTG), digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Source irradiated bowel harvested 72 hours after
irradiation (1400 Gys). Library constructed by Bob
Barstead."
BASE COUNT 110 a 108 c 106 g 115 t
ORIGIN

Alignment Scores:
Pred. No.: 3,39e-05 Length: 439
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.46% Indels: 0
DB: 10 Gaps: 0

US-08-836-455-4 (1-153) x BE136816 (1-439)

QY 103 SerSerLeuthrSerGluAspSerAlaValTyrPheCysAlaArgGly 118
|||||
Db 367 AGCAGCCTAACATCTGAGGACTCTGGGCTCTATTCTGTGCAAGAGGA 414

RESULT 15
CA577954
LOCUS
DEFINITION
CA577954 529 bp mRNA linear EST 19-NOV-2002
K0716F08-5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)
cDNA Library (Long) Mus musculus cDNA clone NIA:K0716F08
IMAGE:30075139 5', mRNA sequence.
CA577954
CA577954.1 GI:25126345
EST.
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 529)
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
Aiba, K., Raub, D., Longo, D.L., Keller, J. and Ko, M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)
Unpublished
Other ESTs: K0716F08-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0716 row: F column: 08
Seq primer: M13 Reverse
High quality sequence stop: 529
POLYA=No.
FEATURES
source
1. .529
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6NCr"
/db_xref="niaEST:K0716F08-5N"
/db_xref="taxon:10090"
/clone="NIA:K0716F08 IMAGE:30075139"
/tissue_type="Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)"
/dev_stage="Age approx.10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)"
/notes="Vector: pSPORT1 (Invitrogen); Site_1: Salt; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,"

```

National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were obtained from Drs. Dennis Taub, Dan Longo (National Institute on Aging, USA), Jonathan Keller (National Cancer Institute, USA). Double-stranded cDNAs were synthesized with an oligo(dT) primer [Invitrogen: 5'-pgactatgttctagatcgcgagcgccgctttttttttttt-3'] from 0.9 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker L1-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.1 kb. The library was constructed by Yulan Piao (NIA)."

BASE COUNT 113 a 146 c 139 g 131 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.07e-05 Length: 529  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.46% Indels: 0  
DB: 14 Gaps: 0

US-08-836-455-4 (1-153) x CA577954 (1-529)

Qy 103 SerSerLeuThrSerGluAspSerAlaValTyrPheCysAlaArgGly 118  
Db 329 AGCAGCCTGACCTCTGAGGACTCTGCGGCTCTATTCTGTGCAAGAGGG 376

Search completed: August 31, 2003, 02:16:28  
Job time : 1845.13 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 19:50:30 ; Search time 36.453 Seconds  
(without alignments)  
666.204 Million cell updates/sec

Title: US-08-836-455-4  
Perfect score: 816  
Sequence: 1 MCSWVFLPILLSITGVHSQ.....TVSSAKTTPPVYPLVPGSL 153

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	816	100.0	153	18	AAW27120 Murine monoclonal
2	816	100.0	153	20	AAW87594 Antibody 11D10 hea
3	816	100.0	133	24	AAO16293 Mouse 11D10 antibo
4	621	76.1	145	16	AAO66758 Anti-tobacco mosai
5	602	73.8	159	10	AAO93079 Heavy chain of mon
6	574.5	70.4	468	12	AAO13061 Monoclonal antibod
7	568	69.6	561	20	AAV17415 Mouse immunoglobul
8	565	69.2	140	10	AAO94780 2 H7 VH gene. Syn
9	565	69.2	140	18	AAW10588 2H7 antibody heavy

10	565	69.2	140	18	AAW16343 2H7 heavy chain va
11	565	69.2	140	18	AAW10242 Heavy chain variab
12	565	69.2	140	19	AAW41070 Mouse 2H7 antibody
13	565	69.2	140	19	AAW47513 Mouse 2H7 antibody
14	565	69.2	140	19	AAW47520 Mouse 2H7 antibody
15	565	69.2	140	20	AAW89540 Mouse antibody 2H7
16	565	69.2	140	22	AAW98092 2H7 heavy chain va
17	563	69.0	463	18	AAW14939 3F4 (Chimeric) hum
18	563	69.0	463	18	AAW14940 3F4 (Chimeric) hum
19	563	69.0	464	18	AAW14941 Murine anti-porcine
20	563	69.0	464	18	AAW14938 Monoclonal antibod
21	563	69.0	469	14	AAW40384 Human CH2 domain d
22	561.5	68.8	360	23	AAE27924 Antibody C2B8 CH2
23	561.5	68.8	360	21	ABB82833 A dimeric anti-CD2
24	561.5	68.8	470	21	AAO80826 Human C2B8 antibod
25	561.5	68.8	470	23	AAE27923 Antibody C2B8 heav
26	561.5	68.8	470	24	ABB82832 Sequence encoded b
27	561	68.8	140	8	AAW70627 Anti-Fas MAb HFE7A
28	560.5	68.7	464	19	AAW83041 Mouse anti-Fas ant
29	560.5	68.7	464	21	AAW14747 Murine anti-Fas an
30	560.5	68.7	464	21	AAW90897 Humanised anti-Fas
31	560.5	68.7	464	23	ABB74866 Humanised anti-Fas
32	560.5	68.7	464	23	ABB74912 Human penton base
33	554.5	68.0	438	23	AAE18372 Portion of mouse D
34	554.5	68.0	438	24	ABG76347 Human penton base
35	554.5	68.0	456	23	AAE18370 Human penton base
36	554.5	68.0	456	24	ABG76345 Mouse DAV-1 heavy
37	554.5	68.0	464	16	AAW76088 MAb 55.1 heavy cha
38	554.5	68.0	493	23	AAE18379 Human N-terminal D
39	554.5	68.0	493	24	ABG76354 Mouse DAV-1 heavy
40	554.5	68.0	510	23	AAE18378 Human N-terminal D
41	554.5	68.0	510	24	ABG76353 Mouse DAV-1 heavy
42	554.5	68.0	597	23	AAE18377 Human N-terminal D
43	554.5	68.0	597	24	ABG76352 Mouse DAV-1 heavy
44	554.5	68.0	613	23	AAE18380 Human N-terminal D
45	554.5	68.0	613	24	ABG76355 Mouse DAV-1 heavy

ALIGNMENTS

RESULT 1  
AAW27120  
ID AAW27120 standard; Protein; 153 AA.  
XX AC AAW27120;  
XX DT 25-MAR-2003 (updated)  
XX DE 04-JAN-1998 (first entry)  
XX Murine monoclonal anti-idiotypic antibody 11D10 VH region.  
XX Monoclonal antibody 11D10; anti-idiotypic antibody; mucin;  
XX human milk fat globule; HMF; tumour; breast cancer; vaccine.  
XX Mus musculus.  
XX Key Location/Qualifiers  
XX Peptide 1..19  
XX Protein /label= Sig\_peptide  
XX Region /label= Mat\_protein  
XX /label= PR1  
XX /note= "framework region 1"  
XX Region 50..54  
XX /label= CDR1  
XX /note= "complementarity determining region 1"  
XX Region 55..68  
XX /label= PR2  
XX /note= "framework region 2"  
XX Region 69..85  
XX /label= CDR2

```

FT      Region      /note= "complementarity determining region 2"
FT      FT      86..117
FT      FT      /label= FR3
FT      Region      /note= "framework region 3"
FT      FT      118..126
FT      FT      /label= CDR3
FT      Region      /note= "complementarity determining region 3"
FT      FT      127..137
FT      FT      /label= FR1
FT      Region      /note= "framework region 4"
FT      FT      138..153
FT      FT      /label= Constant
XX      XX
XX      PN      WO9722699-A2.
XX      PD      26-JUN-1997.
XX      PF      19-DEC-1996; 96WO-US20757.
XX      PR      20-DEC-1995; 95US-0575762.
XX      PR      26-JAN-1996; 96US-0591965.
XX      PR      13-DEC-1996; 96US-0766350.
XX      XX
XX      PA      (KENT ) UNIV KENTUCKY.
XX      XX
XX      PI      Chatterjee M, Chatterjee SK, Foon KA;
XX      XX      WPI; 1997-341690/31.
XX      DR      N-PSDB; AAT85150.
XX      XX
XX      Monoclonal anti-idiotype antibody 11D10 - elicits immune response
XX      PT      against human milk fat globule disease associated tumours,
XX      PT      especially breast cancer.
XX      PS      Claim 10; Page 94-95; 130pp; English.
XX      CC
XX      CC      This polypeptide sequence comprises the heavy chain variable region
XX      CC      (VH) of monoclonal anti-idiotype antibody 11D10 produced by
XX      CC      hybridoma cell line ATCC 12020. 11D10 was obtained by immunising
XX      CC      naive mice with MC-10 anti-HMFG antibody to obtain an anti-idiotype
XX      CC      response. It elicits an immune response against a specific epitope
XX      CC      of a high mol.wt. mucin of human milk fat globule (HMFG). It
XX      CC      induces an immunological response to HMFG in mice, rabbits, monkeys
XX      CC      and patients with advanced HMFG-associated tumours. Pharmaceutical
XX      CC      compositions and vaccines comprising 11D10, 11D10 polypeptides
XX      CC      and/or 11D10 polynucleotides (see also AAT85149-50) are claimed.
XX      CC      Also claimed are diagnostic kits and methods of using 11D10, 11D10
XX      CC      polypeptides and/or 11D10 polynucleotides, including methods of
XX      CC      treating HMFG-associated tumours. 11D10 is also used in a claimed
XX      CC      method of palliating HMFG-associated disease and in claimed kits to
XX      CC      detect or quantify anti-HMFG antibody.
XX      CC      (Updated on 25-MAR-2003 to correct PR field.)
XX      SQ      Sequence 153 AA;
XX      Query Match 100.0%; Score 816; DB 18; Length 153;
XX      Best Local Similarity 100.0%; Pred. No. 3.7e-55;
XX      Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 MECSWFLFLSITGVHQSQAYLQOAGELVRSGASVKMSCKASGYLTITSNMHWKQTP 60
DB      1 MECSWFLFLSITGVHQSQAYLQOAGELVRSGASVKMSCKASGYLTITSNMHWKQTP 60
OY      61 GQGLEWIGNTPPGNDTYNQKFKGASLTADTSSSTAYMQISLTSEDSAVYFCARGNW 120
DB      61 GQGLEWIGNTPPGNDTYNQKFKGASLTADTSSSTAYMQISLTSEDSAVYFCARGNW 120
OY      121 EGALDYGQGSVTVSSAKTTPPPVPLVPGSL 153
DB      121 EGALDYGQGSVTVSSAKTTPPPVPLVPGSL 153

```

RESULT 2

```

AAW87594
ID      AAW87594 standard; Protein; 153 AA.
XX      AC      AAW87594;
XX      DT      16-MAR-1999 (first entry)
XX      DE      Antibody 11D10 heavy chain variable region.
XX      KW      Murine; mouse; antibody; heavy chain; variable region; anti-idiotype;
XX      KW      human milk fat globule; tumour; ovary; lung; pancreas; carcinoma; breast.
XX      OS      Mus sp.
XX      FH      Key
XX      FT      Peptide
XX      FT      1..19
XX      FT      /note= "signal peptide"
XX      FT      Protein
XX      FT      20..153
XX      FT      /note= "mature protein"
XX      FT      Region
XX      FT      20..49
XX      FT      /label= framework_1
XX      FT      Domain
XX      FT      50..54
XX      FT      /label= CDR1
XX      FT      /note= "complementarity determining region 1"
XX      FT      Region
XX      FT      55..68
XX      FT      /label= framework_2
XX      FT      Domain
XX      FT      69..85
XX      FT      /label= CDR2
XX      FT      /note= "complementarity determining region 2"
XX      FT      Region
XX      FT      86..117
XX      FT      /label= framework_3
XX      FT      Domain
XX      FT      118..126
XX      FT      /label= CDR3
XX      FT      /note= "complementarity determining region 3"
XX      FT      Region
XX      FT      127..137
XX      FT      /label= framework_4
XX      XX
XX      PN      WO9856419-A1.
XX      PD      17-DEC-1998.
XX      PF      12-JUN-1998; 98WO-US12250.
XX      PR      11-JUN-1998; 98US-0096244.
XX      PR      13-JUN-1997; 97US-0049540.
XX      PA      (KENT ) UNIV KENTUCKY RES FOUND.
XX      PI      Chatterjee M, Foon KA;
XX      DR      WPI; 1999-060029/05.
XX      DR      N-PSDB; AAV83773.
XX      XX
XX      Delaying development of, or treating, HMFG-associated tumours -
XX      PT      using anti-idiotype antibody 11D10 raised against antibodies to
XX      PT      human milk fat globule protein
XX      PS      Disclosure; Fig 2; 54pp; English.
XX      CC
XX      CC      This sequence represents the heavy chain variable region of the murine
XX      CC      antibody 11D10. This anti-idiotype antibody is used to delay the
XX      CC      development of, or treat, a human milk fat globule (HMFG) associated
XX      CC      tumour in an individual having low tumour burden. The antibody 11D10
XX      CC      is used to prevent the recurrence of HMFG-associated tumours e.g.
XX      CC      ovarian, non-small cell lung and pancreatic carcinoma, especially for
XX      CC      treating breast tumours.
XX      SQ      Sequence 153 AA;
XX      Query Match 100.0%; Score 816; DB 20; Length 153;
XX      Best Local Similarity 100.0%; Pred. No. 3.7e-55;
XX      Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MECSWVFLLSITITGVHSQAYLQOSGAELVRSYGASVKMSCKASGYTLTSTYNMHWKQTP 60  
 Db 1 MECSWVFLLSITITGVHSQAYLQOSGAELVRSYGASVKMSCKASGYTLTSTYNMHWKQTP 60  
 Qy 61 GQGLEWIGNIPFGNGDYYNOKFKGKASLTADTSSSTAYMIOISLTSEDSAVYFCARGNW 120  
 Db 61 GQGLEWIGNIPFGNGDYYNOKFKGKASLTADTSSSTAYMIOISLTSEDSAVYFCARGNW 120  
 Qy 121 EGALDYGQGTSTVTSSAKTTPPPVPLVPGSL 153  
 Db 121 EGALDYGQGTSTVTSSAKTTPPPVPLVPGSL 153

## RESULT 3

AAO16293  
 ID AAO16293 standard; Protein: 153 AA.

AC AAO16293;

DT 20-MAR-2003 (first entry)

XX Mouse 11D10 antibody heavy chain variable region.

XX Mouse; murine; vaccine; tumour; human milk fat globules; HMFG;

KW carcinoembryonic antigen; CEA; 11D10; 3H1; HMFG-associated tumour;

KW CEA-associated tumour; anti-idiotypic antibody.

XX Mus musculus.

XX Key Location/Qualifiers

FT Peptide 1..19

FT /label= signal\_peptide

FT Protein 20..153

FT /note= "Mature murine 11D10 antibody heavy chain variable region"

XX WO200292012-A2.

XX 21-NOV-2002.

XX 17-MAY-2002; 2002WO-US15840.

XX 17-MAY-2001; 2001US-0861294.

XX (KENT ) UNIV KENTUCKY RES FOUND.

XX Chatterjee M, Foon KA;

XX WPI; 2003-129216/12.

XX N-PSDB; AAL51274.

XX Use of anti-idiotypic antibodies for human milk fat globules (HMFG)- or  
 PT carcinoembryonic antigen (CEA)-associated tumor for delaying the  
 PT development of, or treating a HMFG- or CEA-associated tumor (e.g.  
 PT breast tumor) in humans

XX Claim 2; Fig 2; 98pp; English.

XX The invention comprises a method for delaying the development of, or  
 CC treating a tumour that is associated with human milk fat globules (HMFG)  
 CC or carcinoembryonic antigen (CEA). The method of the invention involves  
 CC administering an anti-idiotypic antibody for HMFG (e.g. 11D10), and an  
 CC anti-idiotypic antibody for CEA (e.g. 3H1). The method is useful for  
 CC delaying the development, of or treating HMFG/CEA-associated tumours. The  
 CC present amino acid sequence represents the heavy chain variable region of  
 CC the mouse 11D10 anti-idiotypic antibody.

XX Sequence 153 AA;

Query Match 100.0%; Score 816; DB 24; Length 153;

Best Local Similarity 100.0%; Pred. No. 3.7e-55;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECSWVFLLSITITGVHSQAYLQOSGAELVRSYGASVKMSCKASGYTLTSTYNMHWKQTP 60  
 Db 1 MECSWVFLLSITITGVHSQAYLQOSGAELVRSYGASVKMSCKASGYTLTSTYNMHWKQTP 60  
 Qy 61 GQGLEWIGNIPFGNGDYYNOKFKGKASLTADTSSSTAYMIOISLTSEDSAVYFCARGNW 120  
 Db 61 GQGLEWIGNIPFGNGDYYNOKFKGKASLTADTSSSTAYMIOISLTSEDSAVYFCARGNW 120  
 Qy 121 EGALDYGQGTSTVTSSAKTTPPPVPLVPGSL 153  
 Db 121 EGALDYGQGTSTVTSSAKTTPPPVPLVPGSL 153

## RESULT 4

AAO66758  
 ID AAR66758 standard; Protein: 465 AA.

AC AAR66758;

DT 01-SEP-1995 (first entry)

XX Anti-tobacco mosaic virus monoclonal Ab heavy chain.

XX Tobacco mosaic virus; TMV; monoclonal antibody;

KW heavy chain; virus-resistant plants; biofarming.

XX Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..19

FT /label= leader

FT Peptide 20..465

FT /label= mat\_peptide

FT Domain 20..128

FT /note= "variable heavy domain"

FT Domain 129..141

FT /note= "J heavy 4 domain"

FT Domain 142..465

FT /note= "constant heavy domain"

XX JP06319396-A.

XX 22-NOV-1994.

XX 07-MAY-1993; 93JP-0131208.

XX 07-MAY-1993; 93JP-0131208.

XX (NISR ) JAPAN TOBACCO INC.

XX (KURS ) KURARAY CO LTD.

XX WPI; 1995-040220/06.

XX N-PSDB; AAO79930.

XX Transformed plant producing animal-derived anti-virus antibody -  
 PT esp. tobacco plants producing anti-tobacco mosaic virus  
 PT monoclonal antibody

XX Example 2; Pages 14-15; 26pp; Japanese.

XX AAQ79929 and AAQ79930 encode AAR66757 and AAR66758, the light and heavy  
 CC chains of an animal derived anti-tobacco mosaic virus (TMV)  
 CC monoclonal antibody. The cDNAs were incorporated into a T1  
 CC plasmid vector, which was incorporated into A. tumefaciens.  
 CC The resultant plant expression vector was used to transform  
 CC tobacco plants, making them TMV resistant, the plants could  
 CC also be biofarmed for the prodn. of anti-virus antibodies.

XX Sequence 465 AA;

Query Match 76.1%; Score 621; DB 16; Length 465;

Best Local Similarity 75.6%; Pred. No. 1e-39;

Matches 118; Conservative 15; Mismatches 19; Indels 4; Gaps 2;

```

Qy 1 MECSWFLFLSITGTGVSQAYLQOQSGAELVRSGASVKMSCKASGYTLTSTNMHWKQTP 60
Db 1 MECNWILPFLSVTSYGVSQVQLQOQSGAELARPGASVKLSCKASGYFTSYWQWVKQRP 60
Qy 61 GQGLEWIGNIFPGNGDYYNOKFKGKASLTADTSSSTAYMQISSLTSEDSAVYFCARG-- 118
Db 61 GQGLEWIGAIYPGNGDTRYTKFKGKATLTADKSSSTAYMQLSALASEDSAVYFCAREGG 120
Qy 119 -NW-EGALDYWGQGTSTVTSSAKTTPPPVPLVPGS 152
Db 121 YNSDYANDYWGQGTSTVTSSAKTTPPSVPLAPGS 156

RESULT 5
ID AAP93079 standard; peptide: 159 AA.
XX
AC AAP93079;
XX
DT 25-MAR-2003 (updated)
DT 31-OCT-2002 (updated)
DT 15-MAR-1990 (first entry)
XX
DE Heavy chain of monoclonal antibody 6A4.
XX
KW Monoclonal antibody 6A4; heavy chain; Pseudomonas aeruginosa; OMP-1.
XX
OS Unidentified.
XX
PN EP338395-A.
XX
PD 25-OCT-1989.
XX
PF 12-APR-1989; 89EP-0106463.
XX
PR 19-APR-1988; 88DE-3813023.
XX
PA (BEHW ) BEHRINGER AG.
XX
PI Domdey H, Marget M, Vonspecht B;
XX
WPI: 1989-310861/43.
DR N-PSDB; RAN91645.
XX
Monoclonal antibody to Pseudomonas aeruginosa - and DNA coding for
variable antibody regions.
XX
PS Claim 1; page 6; 7pp; german.
XX
The peptide is encoded by the heavy chain of monoclonal antibody 6A4.
CC 6A4 reacts with the OMP-1 protein of all 19 known serotypes of
CC P.aeruginosa. It is used for therapy and diagnosis of infection, and as
CC a carrier for drugs. The antibody is IgG2a subclass.
CC (Updated on 31-OCT-2002 to add missing OS field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 159 AA;
Query Match 73.8%; Score 602; DB 10; Length 159;
Best Local Similarity 76.3%; Pred. No. 9.4e-39;
Matches 116; Conservative 9; Mismatches 25; Indels 2; Gaps 1;
Qy 1 MECSWFLFLSITGTGVSQAYLQOQSGAELVRSGASVKMSCKASGYTLTSTNMHWKQTP 60
Db 1 MERHWILFLESVTAGVHSQVQLQOQSGAELARPGASVKMSCKASGYFTSYWQWVKQRP 60
Qy 61 GQGLEWIGNIFPGNGDYYNOKFKGKASLTADTSSSTAYMQISSLTSEDSAVYFCARG-- 118
Db 61 GQGLEWIGYINPTNGYTYNQNFKDKATLTADKSSSTAYMQLSLTSEDSAVYCYTRSY 120
Qy 119 NWEGALDYWGQGTSTVTSSAKTTPPPVPLVP 150

```

```

Db 121 NYEGAMDYWGQGTSTVTSSAKTTPAPSVPLAP 152

RESULT 6
ID AAR13061 standard; Protein: 468 AA.
XX
AC AAR13061;
XX
DT 25-MAR-2003 (updated)
DT 03-OCT-1991 (first entry)
XX
DE Monoclonal antibody OK3T heavy chain.
XX
KW OK3T; light chain; humanised antibodies; CDR-grafting.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT Protein /label= signal peptide
FT /label= light chain
XX
PN W09109967-A.
XX
PD 11-JUL-1991.
XX
PF 21-DEC-1990; 90WO-GB02017.
XX
PR 21-DEC-1990; 90WO-GB02017.
XX
PR 21-DEC-1989; 89GB-0028874.
XX
PA (CLLT ) CELLTech LTD.
XX
PI Adair JR, Athwal DS, Emtage JS;
XX
WPI: 1991-222915/30.
DR N-PSDB; AAQ12637.
XX
New humanised antibodies comprising CDR grafted antibody - with
heavy and light chains, for use in vivo therapy and diagnosis
XX
Disclosure: Fig 2b; 91pp; English.
XX
The OK3T heavy chain sequence was deduced from the cDNA sequence
isolated from a library prepared from OK3T producing cells. The
library was screened with a probe complementary to a region in the
mouse IgG2a constant domain region. The OK3T sequence was used in
CDR-grafting experiments to prepare humanised antibodies.
XX
(Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 468 AA;
Query Match 70.4%; Score 574.5; DB 12; Length 468;
Best Local Similarity 74.8%; Pred. No. 3.7e-36;
Matches 113; Conservative 8; Mismatches 29; Indels 1; Gaps 1;
Qy 1 MECSWFLFLSITGTGVSQAYLQOQSGAELVRSGASVKMSCKASGYTLTSTNMHWKQTP 60
Db 1 MERHWILFLLSVTAGVHSQVQLQOQSGAELARPGASVKMSCKASGYFTTRYTMHWVKQRP 60
Qy 61 GQGLEWIGNIFPGNGDYYNOKFKGKASLTADTSSSTAYMQISSLTSEDSAVYFCAR--GN 119
Db 61 GQGLEWIGYINPSRGYTYNQNFKDKATLTADKSSSTAYMQLSLTSEDSAVYCYARYD 120
Qy 120 WEGALDYWGQGTSTVTSSAKTTPPPVPLVP 150
Db 121 DHYCLDYWGQGTSTVTSSAKTTPAPSVPLAP 151

RESULT 7

```





QY 118 -GNWEGALDYWGQGTSTVTVS 136  
 DB 121 YSNSYWFYDVGWGTGTTTVS 140

RESULT 9  
 AAW10588  
 ID AAW10588 standard; Protein: 140 AA.  
 XX AAW10588;  
 XX 25-MAR-2003 (updated)  
 DT 21-OCT-1997 (first entry)  
 XX 2H7 antibody heavy chain variable region.  
 XX Immunoglobulin G; IgG; heavy chain; recombinant production;  
 KW antibody; passive immunisation; serum sickness; anaphylactic shock;  
 KW immunosay; imaging; reagent; complement mediated lysis;  
 KW therapy; variable; region; complementarity determining; CDR;  
 XX framework.

Mus spp.  
 OS  
 XX Key Location/Qualifiers  
 FH Peptide 1..19  
 FT Peptide /label= sig\_peptide  
 FT Peptide 20..140  
 FT Peptide /label= mat\_peptide  
 FT Region 20..49  
 FT Region /label= framework\_1  
 FT Region 50..54  
 FT Region /label= CDR\_1  
 FT Region 55..68  
 FT Region /label= framework\_2  
 FT Region 69..85  
 FT Region /label= CDR\_2  
 FT Region 86..117  
 FT Region /label= framework\_3  
 FT Region 118..130  
 FT Region /label= CDR\_3  
 FT Region 131..140  
 FT Region /label= framework\_4

US5595898-A.  
 XX  
 XX 21-JAN-1997.  
 PD  
 XX 18-AUG-1994; 94US-0299085.  
 DF  
 XX 29-MAR-1990; 90US-0501092.  
 PR 01-NOV-1985; 85US-0793980.  
 PR 27-OCT-1986; 86WO-US02269.  
 PR 24-JUL-1987; 87US-0077528.  
 PR 11-JAN-1988; 88US-0142039.  
 PR 08-DEC-1992; 92US-0987555.  
 PR 18-AUG-1994; 94US-0299085.  
 XX  
 XX (XOMA ) XOMA CORP.  
 XX  
 XX Better M, Horwitz AH, Lei S, Liu AY, Robinson RR;  
 PI Wall R, Wilcox GL;  
 FI  
 XX WPI; 1997-107579/10.  
 XX N-PSDB; AAT36316.  
 XX  
 XX Nucleic acid encoding immunoglobulin fragment - comprising  
 PT di:cistronic transcription unit with pectate lyase signal sequences  
 FT Example; Fig 21; 95pp; English.  
 XX  
 XX The present sequence is the heavy chain variable region of

CC the 2H7 antibody. The 2H7 cDNA was used in the preparation of  
 CC a novel polynucleotide molecule encoding an Ig fragment. The DNA  
 CC molecule comprises 2 DNA sequences encoding 2 pectate lyase  
 CC secretion signal sequences respectively linked to a DNA sequence  
 CC encoding an Ig Fd molecule or Ig light chain, operably linked to a  
 CC single prokaryotic promoter so as to form a dicistronic  
 CC transcription unit, provided that the Ig fragment can bind an  
 CC antigen and is produced and secreted by an E. coli host cell when  
 CC the nucleic acid molecule is expressed in the host cell.  
 CC The polynucleotide molecule is used for the production of  
 CC recombinant antibodies, which can be used for passive immunisation  
 CC without negative immune reactions (e.g. serum sickness and  
 CC anaphylactic shock), in labelled forms as immunoassay or imaging  
 CC reagents, in complement mediated lysis and for therapeutic  
 CC purposes when coupled to a toxin or other therapeutic agent.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 140 AA;  
 Query Match 69.2%; Score 565; DB 18; Length 140;  
 Best Local Similarity 80.0%; Pred. No. 5.6e-36;  
 Matches 112; Conservative 6; Mismatches 18; Indels 4; Gaps 1;  
 QY 1 MECSWVFLFLSLITGVHSGAYLQOCSAELVRSASVYKMSCKASGYTLTSYNMHWVKQTP 60  
 DB 1 MGFSRIPLFLSLVTTGVHSGAYLQOCSAELVRSASVYKMSCKASGYTLTSYNMHWVKQTP 60  
 QY 61 GQGLEWIGNIPPGNGDTYYNOKFKGKASLTADTSSSTAYMQISLTSSEDSAVYFCAR--- 117  
 DB 61 RQGLEWIGAIYPGNGDTSYNOKFKGKATLIVDKSSSTAYMQISLTSSEDSAVYFCARVY 120  
 QY 118 -GNWEGALDYWGQGTSTVTVS 136  
 DB 121 YSNSYWFYDVGWGTGTTTVS 140

RESULT 10  
 AAW16343  
 ID AAW16343 standard; Protein: 140 AA.  
 XX AAW16343;  
 XX 25-MAR-2003 (updated)  
 DT 04-SEP-1997 (first entry)  
 XX 2H7 heavy chain variable sequence.  
 XX Antibody engineering; heavy chain; light chain; chimaeric antibody;  
 KW passive immunisation; diagnosis; hybridoma; monoclonal antibody;  
 KW 2H7; B-cell antigen; Bp35.  
 XX Mus sp.  
 XX Key Location/Qualifiers  
 FH Peptide 1..19  
 FT Peptide /label= sig\_peptide  
 FT Region 20..49  
 FT Region /label= FR1  
 FT Region /note= "framework region 1"  
 FT Region 50..54  
 FT Region /label= CDR1  
 FT Region /note= "complementarity determining region 1"  
 FT Region 55..68  
 FT Region /label= FR2  
 FT Region /note= "framework region 2"  
 FT Region 69..85  
 FT Region /label= CDR2  
 FT Region /note= "complementarity determining region 2"  
 FT Region 86..117  
 FT Region /label= FR3  
 FT Region /note= "framework region 3"  
 FT Region 118..130  
 FT Region /label= CDR3

FT /note= "complementarity determining region 3"  
 FT 131..140  
 FT /label= FR4  
 FT /note= "framework region 4"  
 XX  
 PN US5618920-A.  
 XX  
 XX 08-APR-1997.  
 XX 29-APR-1994; 94US-0235225.  
 XX 29-MAR-1990; 90US-0501092.  
 PR 01-NOV-1985; 85US-0793980.  
 PR 27-OCT-1986; 86WO-US02269.  
 PR 24-JUL-1987; 87US-0077528.  
 PR 11-JAN-1988; 88US-0142039.  
 PR 17-APR-1992; 92US-0870404.  
 PR 29-APR-1994; 94US-0235225.  
 XX  
 PA (XOMA ) XOMA CORP.  
 PI Better M, Horwitz AH, Lei S, Liu AY, Robinson RR;  
 PI Wall R, Wilcox GL;  
 XX WPI; 1997-225473/20.  
 DR N-PSDB; AAT70868.  
 XX  
 XX Secretable immunoglobulin heavy and light chain fragments - capable  
 PT of assembling into chimeric antibodies, useful for e.g. passive  
 PT immunisation, diagnosis, etc  
 XX  
 PS Example 4; Fig 21; 96pp; English.  
 XX  
 CC A polypeptide (AAW16343) comprises the heavy chain variable region of  
 CC mouse monoclonal antibody 2H7, which recognises human B-cell surface  
 CC antigen Bp35. It is encoded by a cDNA clone (AAT70868) obtd. from a  
 CC 2H7 cDNA library by PCR amplification. The 2H7 light chain variable  
 CC sequence (AAW16344) is also provided. A novel human-mouse chimaeric  
 CC antibody with specificity for the human B-cell antigen has been  
 CC constructed.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 XX Sequence 140 AA;  
 SQ  
 Query Match 69.2%; Score 565; DB 18; Length 140;  
 Best Local Similarity 80.0%; Pred. No. 5.6e-36;  
 Matches 112; Conservative 6; Mismatches 18; Indels 4; Gaps 1;  
 QY 1 MECSWFLFLLSITTTGVHSGAYLQOQSGAELVRGASVKMSCKASGYTLTSTNNHWKQTP 60  
 Db 1 MGFSRIFLFLSVTTGVHSGAYLQOQSGAELVRGASVKMSCKASGYTLTSTNNHWKQTP 60  
 QY 61 GQLEWIGNIFPGNGDTYYNQKFKGASLTADTSSSTAYYMQISLTSSEDSAVYFCAR--- 117  
 Db 61 RQLEWIGALYPGNGDTSYNQKFKGATLVLDSSSTAYYMQLSLTSSEDSAVYFCARVY 120  
 QY 118 -GNWEGALDYWGQGTSTVTS 136  
 Db 121 YSNSYWFYDVGWGTGTTVTS 140  
 RESULT 11  
 AAW10242  
 ID AAW10242 standard; Protein; 140 AA.  
 XX  
 AC AAW10242;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 06-AUG-1997 (first entry)  
 XX  
 XX Heavy chain variable sequence of 2H7.  
 XX  
 XX Pectate lyase; signal sequence; Gram-negative bacterium; Immunoglobulin;

KW protein production; human; constant region; passive immunisation; toxin;  
 KW serum sickness; anaphylaxis; sweetener; thaumatin; cytoplasm; periplasm;  
 KW antibody; Ig; heavy-chain; hepatitis; mouse; lung carcinoma; cancer;  
 KW myeloma cell.  
 XX  
 OS Mus musculus.  
 XX  
 XX Key Location/Qualifiers  
 FH Peptide 1..19  
 FT /note= "leader peptide"  
 FT Region 20..49 "framework region 1"  
 FT Region 50..54  
 FT /note= "complementarity determining region 1"  
 FT Region 55..68  
 FT /note= "framework region 2"  
 FT Region 69..85  
 FT /note= "complementarity determining region 2"  
 FT Region 86..117  
 FT /note= "framework region 3"  
 FT Region 118..130  
 FT /note= "complementarity determining region 3"  
 FT Region 131..140  
 FT /note= "framework region 4"  
 XX  
 XX US5576195-A.  
 XX  
 XX 19-NOV-1996.  
 XX  
 PF 09-DEC-1994; 94US-0357234.  
 XX  
 XX 29-MAR-1990; 90US-0501092.  
 PR 01-NOV-1985; 85US-0793980.  
 PR 27-OCT-1986; 86WO-US02269.  
 PR 24-JUL-1987; 87US-0077528.  
 PR 11-JAN-1988; 88US-0142039.  
 PR 08-DEC-1992; 92US-0987555.  
 PR 22-FEB-1993; 93US-0020671.  
 PR 09-DEC-1994; 94US-0357234.  
 XX  
 PA (XOMA ) XOMA CORP.  
 XX  
 XX Better M, Lei S, Robinson RR, Wilcox GL;  
 XX WPI; 1997-011254/01.  
 DR N-PSDB; AAT51042.  
 XX  
 PT Improved prodn. of protein in Gram -ve bacteria using signal  
 PT sequence - from pectate lyase to ensure transport of protein from  
 XX the cytoplasm, esp. for prodn. of antibodies  
 PS Example 4; Fig 21; 86pp; English.  
 XX  
 CC This sequence represents the heavy chain variable region of the 2H7  
 CC mouse monoclonal antibody. The 2H7 antibody recognises the human B-cell  
 CC surface antigen Bp35, which plays a role in B-cell activation. This  
 CC sequence was used in a human-mouse chimeric antibody with human B-cell  
 CC antigen specificity, that was produced using the method of the invention.  
 CC The method of the invention is for the production of a protein in a  
 CC Gram-negative bacterium. The method improves on current techniques, by  
 CC using a vector including DNA encoding the pectate lyase signal sequence  
 CC (see AAT51034), attached to the sequence encoding the protein for  
 CC production. The method is especially used to make immunoglobulins (Ig),  
 CC particularly those with a human constant region, suitable for passive  
 CC immunisation (without risk of serum sickness or anaphylaxis) or for in  
 CC vivo/in vitro diagnosis and imaging. The Ig may also be used  
 CC therapeutically, optionally coupled to toxins, etc. Alternatively the  
 CC protein to be produced is the sweetener thaumatin. The presence of this  
 CC signal sequence means that the protein is exported from the cytoplasm  
 CC and can be recovered from the culture medium or periplasm, in active and  
 CC correctly folded form. The method allows the class of any antibody to be  
 CC switched, e.g. most human-human Ig are of M class, easily reduced and  
 CC aggregated, and these can now be changed to G, A or E classes.

CC (Updated on 25-MAR-2003 to correct PF field.)

SQ Sequence 140 AA;

Query Match 69.2%; Score 565; DB 18; Length 140;  
Best Local Similarity 80.0%; Pred. No. 5.6e-36;  
Matches 112; Conservative 6; Mismatches 18; Indels 4; Gaps 1;

Qy 1 MECSWVFLFLSLITGVHSGAYLQOQSGAELVRSGASVKMSCKASGYTLTSYNNHWWKQTP 60

Db 1 MGFSRIFLFLSLITGVHSGAYLQOQSGAELVRSGASVKMSCKASGYTLTSYNNHWWKQTP 60

Qy 61 GQGLEWIGNIFPGNGDTYYNQKFKGKASLTADTSSSTAYMQISLTSSEDSAVYFCAR--- 117

Db 61 GQGLEWIGNIFPGNGDTYYNQKFKGKASLTADTSSSTAYMQISLTSSEDSAVYFCAR--- 117

Qy 118 -GNWEGALDYWGQTSVTVS 136

Db 121 YSNSYWFYDVGTTVTVS 140

# RESULT 12

AAW41070

ID AAW41070 standard; Protein; 140 AA.

XX AC AAW41070;

XX DT 01-JUN-1998 (first entry)

XX DE Mouse 2H7 antibody heavy chain variable region.

XX KW Mouse; murine; heavy chain; variable region;  
XX KW immunoglobulin fragment production; Ig fragment production;  
XX KW monoclonal antibody 2H7; human B-cell surface antigen.

XX OS Mus sp.

XX FH Key Location/Qualifiers

FT Peptide 1..19

FT Peptide /label= sig\_peptide

FT Peptide 20..140

FT Peptide /label= mat\_peptide

FN US5693493-A.

XX PD 02-DEC-1997.

XX PF 25-MAY-1995; 95US-0450731.

XX PR 29-MAR-1990; 90US-0501092.

XX PR 01-NOV-1985; 85US-0793980.

XX PR 27-OCT-1986; 86WO-US02269.

XX PR 24-JUL-1987; 87US-0077528.

XX PR 11-JAN-1988; 88US-0142039.

XX PR 08-DEC-1992; 92US-0987555.

XX PR 18-AUG-1994; 94US-0239085.

XX PR 25-MAY-1995; 95US-0450731.

XX FA (XOMA ) XOMA CORP.

XX PI Better M, Horwitz AH, Lei S, Liu AY, Robinson RR;

XX PI Wall R, Wilcox GL;

XX XX WPI; 1998-031749/03.

XX DR N-ESDB; AAV03926.

XX PT Production of chimeric antibody fragments - by culturing E. coli

XX PT transformed with dicistronic expression cassette

XX PS Example IV; Fig 21; 98pp; English.

XX CC The present sequence was used in the development of a novel method

XX CC for the production of an immunoglobulin (Ig) fragment capable of

CC binding an antigen. The method comprises culturing an E. coli host  
CC that has been transformed with a nucleic acid molecule encoding the  
CC Ig fragment, under conditions so that the Ig fragment is produced  
CC and secreted. The nucleic acid molecule comprises DNA sequences  
CC encoding: (a) pectate lyase secretion signal sequence operably  
CC linked to a DNA sequence encoding at least the variable region of  
CC an Ig Fd molecule; and (b) pectate lyase secretion signal sequence  
CC operably linked to a DNA sequence encoding at least the variable  
CC region of an Ig light chain, where (a) and (b) are operably linked  
CC to a single prokaryotic promoter to form a dicistronic  
CC transcription unit. The method is used to produce chimeric Fab  
CC molecules, e.g. derived from murine monoclonal antibody 2H7 raised  
CC against human B-cell surface antigen. The invention provides a  
CC novel approach for producing genetically engineered antibodies of  
CC desired variable region specificity and constant region  
CC properties. The cloned Ig gene products can be produced by  
CC expression in genetically engineered organisms. The application of  
CC chemical gene synthesis, recombinant DNA cloning and production of  
CC specific Ig chains in various organisms provides an effective  
CC solution for the efficient large scale production of human  
CC monoclonal antibodies. The invention also provides a solution to  
CC the problem of class switching antibody molecules.

XX SQ Sequence 140 AA;

Query Match 69.2%; Score 565; DB 19; Length 140;

Best Local Similarity 80.0%; Pred. No. 5.6e-36;

Matches 112; Conservative 6; Mismatches 18; Indels 4; Gaps 1;

Qy 1 MECSWVFLFLSLITGVHSGAYLQOQSGAELVRSGASVKMSCKASGYTLTSYNNHWWKQTP 60

Db 1 MGFSRIFLFLSLITGVHSGAYLQOQSGAELVRSGASVKMSCKASGYTLTSYNNHWWKQTP 60

Qy 61 GQGLEWIGNIFPGNGDTYYNQKFKGKASLTADTSSSTAYMQISLTSSEDSAVYFCAR--- 117

Db 61 GQGLEWIGNIFPGNGDTYYNQKFKGKASLTADTSSSTAYMQISLTSSEDSAVYFCAR--- 117

Qy 118 -GNWEGALDYWGQTSVTVS 136

Db 121 YSNSYWFYDVGTTVTVS 140

# RESULT 13

AAW47513

ID AAW47513 standard; Protein; 140 AA.

XX AC AAW47513;

XX DT 05-JUN-1998 (first entry)

XX DE Mouse 2H7 antibody heavy chain variable region.

XX KW Mouse; murine; heavy chain; variable region;  
XX KW immunoglobulin fragment production; Ig fragment production;  
XX KW monoclonal antibody 2H7; human B-cell surface antigen.

XX OS Mus sp.

XX FH Key Location/Qualifiers

FT Peptide 1..19

FT Peptide /label= sig\_peptide

FT Peptide 20..140

FT Peptide /label= mat\_peptide

XX XX US5698435-A.

XX PD 16-DEC-1997.

XX PF 06-JUN-1995; 95US-0467140.

XX XX 29-MAR-1990; 90US-0501092.

XX PR 01-NOV-1985; 85US-0793980.

XX PR 27-OCT-1986; 86WO-US02269.

```

PR 24-JUL-1987; 87US-0077528.
PR 11-JAN-1988; 88US-0142039.
PR 08-DEC-1992; 92US-0987555.
PR 18-AUG-1994; 94US-0299085.
PR 06-JUN-1995; 95US-0467140.
XX
XX (XOMA ) XOMA CORP.
XX
XX Better M, Horwitz AH, Lei S, Liu AY, Robinson RR;
PI Wall R, Wilcox GL;
XX
XX WPI: 1998-051492/05.
DR N-PSDB; AAV18557.
XX
XX DNA encoding secrettable immunoglobulin fragments - comprising at
PT least the variable regions of light or heavy chains
XX
XX Example IV; Fig 21; 98pp; English.
XX
XX The present sequence was used in the development of a novel method
CC for the production of an immunoglobulin (Ig) fragment capable of
CC binding an antigen. The method comprises culturing an E. coli host
CC that has been transformed with a nucleic acid molecule encoding the
CC Ig fragment, under conditions so that the Ig fragment is produced
CC and secreted. The nucleic acid molecule comprises DNA sequences
CC encoding: (a) pectate lyase secretion signal sequence operably
CC linked to a DNA sequence encoding at least the variable region of
CC an Ig Fd molecule; and (b) pectate lyase secretion signal sequence
CC operably linked to a DNA sequence encoding at least the variable
CC region of an Ig light chain, where (a) and (b) are operably linked
CC to a single prokaryotic promoter to form a dicistronic
CC transcription unit. The method is used to produce chimeric Fab
CC molecules, e.g. derived from murine monoclonal antibody 2H7 raised
CC against human B-cell surface antigen. The invention provides a
CC novel approach for producing genetically engineered antibodies of
CC desired variable region specificity and constant region
CC properties. The cloned Ig gene products can be produced by
CC expression in genetically engineered organisms. The application of
CC chemical gene synthesis, recombinant DNA cloning and production of
CC specific Ig chains in various organisms provides an effective
CC solution for the efficient large scale production of human
CC monoclonal antibodies. The invention also provides a solution to
CC the problem of class switching antibody molecules.
XX
XX Sequence 140 AA;
XX
XX Query Match 69.2%; Score 565; DB 19; Length 140;
XX Best Local Similarity 80.0%; Pred. No. 5.6e-36;
XX Matches 112; Conservative 6; Mismatches 18; Indels 4; Gaps 1;
XX
XX QY 1 MECSWVFLFLSITTTGVHSQAYLQSGAELVRSGASVKMSCKASGYTLTSYNNHWVKQTP 60
XX | : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 1 MGFSRIFLLSVTTGVHSQAYLQSGAELVRGASVKMSCKASGYTLTSYNNHWVKQTP 60
XX
XX QY 61 GQGLEWIGNIFPGNGDTYNNQKFKASLTADTSSSTAYMQISLTSEDSAVFYCAR--- 117
XX |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 61 RGLEWIGAYIPGNGDTYNNQKFKATLTVDKSSSTAYMQISLTSEDSAVFYCARVY 120
XX
XX QY 118 -GNWEGALDYWGQTSVTVS 136
XX | : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 121 YSNWYFDWVGRTTVTVS 140
XX
XX RESULT 14
XX AAW47520
XX ID AAW47520 standard; Protein; 140 AA.
XX
XX AC AAW47520;
XX
XX DT
XX XX
XX 05-JUN-1998 (first entry)
XX
XX DE Mouse 2H7 antibody heavy chain variable region.
XX

```

```

KW Mouse; murine; heavy chain; variable region;
KW immunoglobulin fragment production; Ig fragment production;
KW monoclonal antibody 2H7; human B-cell surface antigen.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
FH Peptide 1..19
FT /label= sig_peptide
FT Peptide 20..140
FT /label= mat_peptide
XX
XX US5698417-A.
XX
XX 16-DEC-1997.
XX
XX 06-JUN-1995; 95US-0466203.
XX
XX 29-MAR-1990; 90US-0501092.
PR 01-NOV-1985; 85US-0793980.
PR 27-OCT-1986; 86WO-US02269.
PR 24-JUL-1987; 87US-0077528.
PR 11-JAN-1988; 88US-0142039.
PR 08-DEC-1992; 92US-0987555.
PR 18-AUG-1994; 94US-0299085.
PR 25-MAY-1995; 95US-0450731.
PR 06-JUN-1995; 95US-0466203.
XX
XX (XOMA ) XOMA CORP.
XX
XX Better M, Horwitz AH, Lei S, Liu AY, Robinson RR;
PI Wall R, Wilcox GL;
XX
XX WPI: 1998-051487/05.
DR N-PSDB; AAV18593.
XX
XX Production of recombinant immunoglobulin fragment - comprising Fd
XX molecule and light chain
XX
XX Example IV; Fig 21; 98pp; English.
XX
XX The present sequence was used in the development of a novel method
CC for the production of an immunoglobulin (Ig) fragment capable of
CC binding an antigen. The method comprises culturing an E. coli host
CC that has been transformed with a nucleic acid molecule encoding the
CC Ig fragment, under conditions so that the Ig fragment is produced
CC and secreted. The nucleic acid molecule comprises DNA sequences
CC encoding: (a) pectate lyase secretion signal sequence operably
CC linked to a DNA sequence encoding at least the variable region of
CC an Ig Fd molecule; and (b) pectate lyase secretion signal sequence
CC operably linked to a DNA sequence encoding at least the variable
CC region of an Ig light chain, where (a) and (b) are operably linked
CC to a single prokaryotic promoter to form a dicistronic
CC transcription unit. The method is used to produce chimeric Fab
CC molecules, e.g. derived from murine monoclonal antibody 2H7 raised
CC against human B-cell surface antigen. The invention provides a
CC novel approach for producing genetically engineered antibodies of
CC desired variable region specificity and constant region
CC properties. The cloned Ig gene products can be produced by
CC expression in genetically engineered organisms. The application of
CC chemical gene synthesis, recombinant DNA cloning and production of
CC specific Ig chains in various organisms provides an effective
CC solution for the efficient large scale production of human
CC monoclonal antibodies. The invention also provides a solution to
CC the problem of class switching antibody molecules.
XX
XX Sequence 140 AA;
XX
XX Query Match 69.2%; Score 565; DB 19; Length 140;
XX Best Local Similarity 80.0%; Pred. No. 5.6e-36;
XX Matches 112; Conservative 6; Mismatches 18; Indels 4; Gaps 1;
XX
XX QY 1 MECSWVFLFLSITTTGVHSQAYLQSGAELVRSGASVKMSCKASGYTLTSYNNHWVKQTP 60

```

Db 1 MGFSRIFLLSVTTGVHSGAYLQQSGAELVRPGASVKMSCKASGYTFTSYNMHWKQTP 60  
QY 61 GQGLEWIGNIFPGNGDYYNQKFKGASLTADTSSSTAYMQISSLTSEDSAVYFCAR--- 117  
Db 61 RQGLEWIGAIYPGNGDTCYNQKFKGATLVDKSSSTAYMQISSLTSEDSAVYFCARVY 120  
QY 118 -GNWEGALDYWGQGTSTVTS 136  
Db 121 YSNSYWFYDVGTTGTTVTS 140

## RESULT 15

AAW89540  
ID AAW89540 standard; Protein; 140 AA.

XX AC AAW89540;

XX 20-MAR-2003 (updated)

DT 30-MAR-1999 (first entry)

XX Mouse antibody 2H7 heavy chain variable region.

XX pelB pectate lyase; secretion signal; chimeric antibody;

KW heavy chain; B-cell antigen; antibody 2H7.

XX OS Mus sp.

XX US5846818-A.

XX 08-DEC-1998.

XX 06-JUN-1995; 95US-0472696.

XX 08-DEC-1992; 92US-0987555.

PR 29-MAR-1990; 90US-0501092.

PR 22-FEB-1993; 93US-0020671.

PR 09-DEC-1994; 94US-0357234.

PR 01-NOV-1985; 85US-0793980.

PR 27-OCT-1986; 86WO-US02269.

PR 24-JUL-1987; 87US-0077528.

PR 11-JAN-1988; 88US-0142039.

XX (XOMA ) XOMA CORP.

XX Better M, Horowitz AH, Lei S, Liu AY, Robinson RR;

PI Wall R, Wilcox GL;

XX WPI; 1999-059072/05.

DR N-PSDB; AAW82357.

XX pelB pectate lyase signal sequence - and vector for expression of  
secreted proteins in Gram-negative bacteria

XX Example 4; Fig 21; 98pp; English.

XX The present sequence represents the mouse antibody 2H7 heavy chain  
CC variable region. Antibody 2H7 is specific for human B-cell antigen. The  
CC antibody sequence was used to construct a chimeric human-mouse antibody,  
CC in the course of the invention. The chimeric antibody is expressed in a  
CC secretion vector comprising a pelB pectate lyase secretion signal  
CC peptide. The pelB pectate lyase secretion sequence is useful for  
CC producing a protein such as a chimeric antibody in a bacterial host.  
CC (Updated on 20-MAR-2003 to correct PR field.)

XX Sequence 140 AA;

Query Match 69.28; Score 565; DB 20; Length 140;

Best Local Similarity 80.08; Pred. No. 5.6e-36;

Matches 112; Conservative 6; Mismatches 18; Indels 4; Gaps 1;

QY 1 MECSWVFLFLSVTTGVHSGAYLQQSGAELVRPGASVKMSCKASGYTFTSYNMHWKQTP 60

Db 1 MGFSRIFLLSVTTGVHSGAYLQQSGAELVRPGASVKMSCKASGYTFTSYNMHWKQTP 60  
QY 61 GQGLEWIGNIFPGNGDYYNQKFKGASLTADTSSSTAYMQISSLTSEDSAVYFCAR--- 117  
Db 61 RQGLEWIGAIYPGNGDTCYNQKFKGATLVDKSSSTAYMQISSLTSEDSAVYFCARVY 120  
QY 118 -GNWEGALDYWGQGTSTVTS 136  
Db 121 YSNSYWFYDVGTTGTTVTS 140

Search completed: August 30, 2003, 22:04:05

Job time : 37.453 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	816	100.0	153	9	US-09-861-294-4	Sequence 4, Appl
2	816	100.0	153	12	US-10-367-506-4	Sequence 4, Appl
3	574.5	70.4	468	11	US-09-795-515-7	Sequence 7, Appl
4	560.5	68.7	464	15	US-10-216-484-9	Sequence 9, Appl
5	554.5	68.0	438	10	US-09-503-327A-6	Sequence 6, Appl
6	554.5	68.0	456	10	US-09-503-327A-2	Sequence 2, Appl
7	554.5	68.0	493	10	US-09-503-327A-13	Sequence 13, Appl
8	554.5	68.0	510	10	US-09-503-327A-12	Sequence 12, Appl
9	554.5	68.0	597	10	US-09-503-327A-11	Sequence 11, Appl
10	554.5	68.0	613	10	US-09-503-327A-14	Sequence 14, Appl
11	552.5	67.7	136	15	US-10-160-232-88	Sequence 88, Appl
12	540	66.2	135	12	US-10-244-821-91	Sequence 91, Appl
13	539.5	66.1	180	9	US-09-748-960-6	Sequence 6, Appl
14	530	65.0	139	11	US-09-269-921-105	Sequence 105, Appl
15	530	65.0	139	11	US-09-509-098-4	Sequence 4, Appl

```

; SEQ ID NO 4
; LENGTH: 153
; TYPE: PRT
; ORGANISM: M
; FEATURE:
; NAME/KEY: S
; LOCATION: (
;
US-09-861-294-4

```

QY 61 GQGLEWIGNIPFGNGDYYYNQKFKGKASLTADTSSSTAYMQISSITSEDSAVYFCARGNW 120  
 |||||  
 Db 61 GQGLEWIGNIPFGNGDYYYNQKFKGKASLTADTSSSTAYMQISSITSEDSAVYFCARGNW 120  
 |||||

```
QY 121 EGALDYGQGTSTVTSSAKTTPPVYPLVPGSL 153
DB 121 EGALDYGQGTSTVTSSAKTTPPVYPLVPGSL 153

RESULT 2
US-10-367-506-4
; Sequence 4, Application US/10367506
; Publication No. US20030152575A1
; GENERAL INFORMATION:
; APPLICANT: Malaya CHATTERJEE
; APPLICANT: Kenneth A. FOON
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: TUMORS BEARING HMGF AND CEA ANTIGENS
; FILE REFERENCE: 304142000620
; CURRENT APPLICATION NUMBER: US/10367,506
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US/09/861,294
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/049,540
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 09/095,244
; PRIOR FILING DATE: 1998-06-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(19)
US-10-367-506-4

Query Match 100.0%; Score 816; DB 12; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.4e-62;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECSWVFLSLITGVHSGAYLQOQSGAELVRSGASVKMSCKASGYTLTSTNMHWKQTP 60
DB 1 MECSWVFLSLITGVHSGAYLQOQSGAELVRSGASVKMSCKASGYTLTSTNMHWKQTP 60
QY 61 GQGLEWIGNIPGNGDYNNQKFKGASLTADTSSSTAYMQISLTSEDSAVYFCARGNW 120
DB 61 GQGLEWIGNIPGNGDYNNQKFKGASLTADTSSSTAYMQISLTSEDSAVYFCARGNW 120
QY 121 EGALDYGQGTSTVTSSAKTTPPVYPLVPGSL 153
DB 121 EGALDYGQGTSTVTSSAKTTPPVYPLVPGSL 153

RESULT 3
US-09-795-515-7
; Sequence 7, Application US/09795515
; Publication No. US20030039645A1
; GENERAL INFORMATION:
; APPLICANT: Adair, John R.
; APPLICANT: Athwal, Diljeet S.
; APPLICANT: Emtage, John S.
; TITLE OF INVENTION: Humanised Antibodies
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. US20030039645A1ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/795,515
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/846,658
FILING DATE: 01-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CARP-0057
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-795-515-7

Query Match 70.4%; Score 574.5; DB 11; Length 468;
Best Local Similarity 74.8%; Pred. No. 3.2e-41;
Matches 113; Conservative 8; Mismatches 29; Indels 1; Gaps 1;

QY 1 MECSWVFLSLITGVHSGAYLQOQSGAELVRSGASVKMSCKASGYTLTSTNMHWKQTP 60
DB 1 MECSWVFLSLITGVHSGAYLQOQSGAELVRSGASVKMSCKASGYTLTSTNMHWKQTP 60
QY 61 GQGLEWIGNIPGNGDYNNQKFKGASLTADTSSSTAYMQISLTSEDSAVYFCARGNW 119
DB 61 GQGLEWIGNIPGNGDYNNQKFKGASLTADTSSSTAYMQISLTSEDSAVYFCARGNW 120
QY 120 WEGALDYGQGTSTVTSSAKTTPPVYPLV 150
DB 121 DHYCLDYGQGTSTVTSSAKTTPPVYPLAP 151

RESULT 4
US-10-216-484-9
; Sequence 9, Application US/10216484
; Publication No. US20030103976A1
; GENERAL INFORMATION:
; APPLICANT: Serizawa, No. US20030103976A1ufusa
; APPLICANT: Haruyama, Hideyuki
; APPLICANT: Nakahara, Kaori
; APPLICANT: Tamaki, Ikuko
; APPLICANT: Takahashi, Tohru
; TITLE OF INVENTION: Anti-Fas Antibodies
; FILE REFERENCE: 980126CIP/HG
; CURRENT APPLICATION NUMBER: US/10/216,484
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US/09/499,662
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: US 09/053,583
; PRIOR FILING DATE: 1998-04-01
; NUMBER OF SEQ ID NOS: 165
; SEQ ID NO 9
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-216-484-9

Query Match 68.7%; Score 560.5; DB 15; Length 464;
Best Local Similarity 71.3%; Pred. No. 4.9e-40;
Matches 112; Conservative 10; Mismatches 28; Indels 7; Gaps 2;

QY 1 MECSWVFLSLITGVHSGAYLQOQSGAELVRSGASVKMSCKASGYTLTSTNMHWKQTP 60
DB 1 MGWSCIILFLVATATGVHSGAYLQOQSGAELVRSGASVKMSCKASGYTLTSTNMHWKQTP 60
```



```

RESULT 6
US-09-903-327A-2
; Sequence 2, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erqiang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 456

```

Query Match	68.0%;	Score	554.5;	DB	10;	Length	493;
Best Local Similarity	69.7%;	Pred.	No. 1.7e-39;				
Matches	106;	Conservative	16;	Mismatches	25;	Indels	5;
Gaps	1;						
Qy	1	MECSWFLFLSTTTGVHSCAYLQOQCAEIVRSGASVKMCKASGYTLTSTNNHVVKOTP	60				
Db	1	MGWSWIFLFLSTGTAGVHSEVQOQSGPELVKPGASVKICKASGYTFTDNNHVVKOSH	60				
Qy	61	GGGLEIGNLTFPGNDPTYYNQKFKGASLTADTSSSTAYMQISLTSDEDSAYVFCARGN	120				
Db	61	GKSLEIGIYIYPYKGGTYNQKFKSKATLTDDSSNTAYMELRSLTSDASAVVYCARG--	118				
Qy	121	EGALDWWGQGTSTVTSSAKTTPPPVPLVPGS	152				
Db	119	---IAYWGCGTLVTSSAAKTPPSVYPLVPGS	147				

RESULT 8  
US-09-903-327A-12  
; Sequence 12, Application US/09903327A  
; Patent No. US20020164333A1  
; GENERAL INFORMATION:

```

; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erqiang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein with N-terminal portion of DAV-1 heavy chain
; OTHER INFORMATION: and IGF-1 mature peptide
US-09-903-327A-12

Query Match      58.0%; Score 554.5; DB 10; Length 510;
Best Local Similarity 69.7%; Pred. No. 1.1e-39;
Matches 106; Conservative 16; Mismatches 25; Indels 5; Gaps 1;

QY      1 MECSWFLFLLSTTTGVHSGAYLQQSGAEILVRGSAVKMSCKASGYTLTSYNMHWKQTP 60
Db      1 MGWSWFLFLLSTAGVHSEVQLQQSGPELVKFGASVKISCKASGYTFTDYNMHWKQSH 60
          | |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      61 GGLEWIGNIFFPGNGDTYYNQKFKGASLTADTSSSTAYMQISLTSSEDSAVYFCARGNW 120
Db      61 GKSLWTGIYYPKGGTYNQKFKSKATLTDDSSNTAYMELRSLTSDASAVYVCARG-- 118
          | ||||| :| :| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      121 EGALDWGQGTSVTVSSAKTTPPPVPYLPVGS 152
Db      119 ---IAYWGQGLTVTVSAAKTTPPSVYPLAFGS 147
          | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-09-903-327A-11
; Sequence 11, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erqiang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein with N-terminal portion of DAV-1 heavy chain
; OTHER INFORMATION: and TNF alpha mature peptide
US-09-903-327A-11

Query Match      68.0%; Score 554.5; DB 10; Length 597;
Best Local Similarity 69.7%; Pred. No. 2.1e-39;
Matches 106; Conservative 16; Mismatches 25; Indels 5; Gaps 1;

QY      1 MECSWFLFLLSTTTGVHSGAYLQQSGAEILVRGSAVKMSCKASGYTLTSYNMHWKQTP 60
Db      1 MGWSWFLFLLSTAGVHSEVQLQQSGPELVKFGASVKISCKASGYTFTDYNMHWKQSH 60
          | |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      61 GGLEWIGNIFFPGNGDTYYNQKFKGASLTADTSSSTAYMQISLTSSEDSAVYFCARGNW 120
```

```

Db      61 GKSLEWIGIYIPKGGTGYNQKEKSKATLTDTSSNTAYMELRSLTSDASAVYYCARG-- 118
        :|::||| |::| | |||||| |::| |::| |::| |::| |::| |::| |::| |::|
Qy     121 EGALDWGQGTSVTVSSAKTTTPPVYPVLPGS 152
        :|::||| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db     119 ---IAYWGQGLTVLSAAKTTPPSVYPLAPGS 147

RESULT 10
US-09-903-327A-14
; Sequence 14, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erguang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARTARIC ACID
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein with N-terminal portion of DAV-1 heavy chain
; OTHER INFORMATION: and SCF mature peptide
US-09-903-327A-14

Query Match          68.0%; Score 554.5; DB 10; Length 613;
Best Local Similarity 69.7%; Pred. No. 2,2e-39;
Matches 106; Conservative 16; Mismatches 25; Indels 5; Gaps 1;

Qy       1 MECSWFFLFTLITGVHSGAQYLQQSGAEILVRSGASKVMKSCAGTYTLTSYNMHVVKQT 60
         :|::||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db       1 MGHSWFLLFLLSTAGVSESVQLQQSGPELVKFGASVKISCKASGYTFDYNMHVWKQS 60
         :|::||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Qy       61 GGQLEWTGNIFPONGDTYYNKKFKGRKASLPADTSSSTAYMQISLTSDEDSAVTYFCARGNW 120
         :|::||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db       61 GKSLIEWIGIYIPKGGTGYNQKEKSKATLTDTSSNTAYMELRSLTSDASAVYYCARG-- 118
        :|::||| |::| | |||||| |::| |::| |::| |::| |::| |::| |::| |::|
Qy     121 EGALDWGQGTSVTVSSAKTTTPPVYPVLPGS 152
        :|::||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db     119 ---IAYWGQGLTVLSAAKTTPPSVYPLAPGS 147


RESULT 11
US-10-160-232-88
; Sequence 88, Application US/10160232
; Publication No. US20030088075A1
; GENERAL INFORMATION:
; APPLICANT: SHITARA, KENYA
; APPLICANT: ITO, MIKIYO
; APPLICANT: HANAI, NOBUO
; APPLICANT: KAWADA, YOKO
; APPLICANT: NAKAMURA, KAZUYASU
; APPLICANT: SHIBUYA, MASABUMI
; TITLE OF INVENTION: ANTI-HUMAN VEGF RECEPTOR FLT-1 MONOCLONAL ANTIBODY
; FILE REFERENCE: 249-107
; CURRENT APPLICATION NUMBER: US/10/160,232
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US/09/453,718
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 09/315,051
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/119,014
; PRIOR FILING DATE: 1998-07-20
; PRIOR APPLICATION NUMBER: PCT/Jp97/04259
```

```

; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erqiang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein with N-terminal portion of DAV-1 heavy chain
; OTHER INFORMATION: and IGF-1 mature peptide
US-09-903-327A-12

Query Match      58.0%; Score 554.5; DB 10; Length 510;
Best Local Similarity 69.7%; Pred. No. 1.1e-39;
Matches 106; Conservative 16; Mismatches 25; Indels 5; Gaps 1;

QY      1 MECSWFLFLLSTTTGVHSGAYLQQSGAEILVRGASVKMSCKASGYTLTSYNNHHWKQTP 60
        | |||:||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db      1 MGWSWFLFLLSTAGVHSEVQLQQSGPELVKPGASVKISCKASGYTFDTYNHHWKQSH 60
        | |||:||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

QY      61 GGLEWIGNIFFPNGDITYYNQKFKGASLTADTSSSTAYMQISLTSSEDSAVYFCARGNW 120
        | |||| | : | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db      61 GKSLWTGIYYPKGGTYNQKFKSKATLTDDSSNTAYMELRSLTSDASAVYVCARG-- 118
        | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

QY      121 EGALDWGQGTSVTVSSAKTTPPPVPLVPGS 152
        | ||||| ||||| ||||| ||||| |||| |
Db      119 ---IAYWGQGLTVTVSAKTTPPSVYPLAFGS 147
        | ||||| ||||| ||||| ||||| |||| |

RESULT 9
US-09-903-327A-11
; Sequence 11, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erqiang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein with N-terminal portion of DAV-1 heavy chain
; OTHER INFORMATION: and TNF alpha mature peptide
US-09-903-327A-11

Query Match      68.0%; Score 554.5; DB 10; Length 597;
Best Local Similarity 69.7%; Pred. No. 2.1e-39;
Matches 106; Conservative 16; Mismatches 25; Indels 5; Gaps 1;

QY      1 MECSWFLFLLSTTTGVHSGAYLQQSGAEILVRGASVKMSCKASGYTLTSYNNHHWKQTP 60
        | |||:||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db      1 MGWSWFLFLLSTAGVHSEVQLQQSGPELVKPGASVKISCKASGYTFDTYNHHWKQSH 60
        | |||:||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

QY      61 GGLEWIGNIFFPNGDITYYNQKFKGASLTADTSSSTAYMQISLTSSEDSAVYFCARGNW 120

```

```

; PRIOR FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 88
; LENGTH: 136
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; -OTHER INFORMATION: Description
US-160-232-88

```

[illegible]

RESULT 12

```

US-10-244-821-91
; Sequence 91, Application US/10244821
; Publication No. US2003014323A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EX
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/2
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Versio
; SEQ ID NO 91
; LENGTH: 135
; TYPE: PR
; ORGANISM: Mus musculus
US-10-244-821-91

```

Query Match	66.2%	Score 540;	DB 12;	Length 135;
Best Local Similarity	75.9%;	Pred. No. 7.1e-39;		
Matches 104;	Conservative 8;	Mismatches 23;	Indels 2;	Gaps 1;
QY	1	MECSWFLFLSLTTGVHSGAYLQOQSGAEALVRSVASVKMSCKASGYTLTSYNMHHWVKQP	60	
Db	1	MERHWFLFLSVTAGVHSGVQVQOQSGAEALAKPGASVKMSCKASGYTFTSYRMHHWVKRP	60	
QY	61	GCGLEWIGNTFPQNGDITYYNQKFKGASLTADTSSSTAYMQISLTSSEDSAVYFCARGNW	120	
Db	61	GCGLEWIGLYNPSTGYTEYNQKFKDKATLTADSSSTAYMQLSLTFEDSAVYFCARGG	119	
QY	121	EGALDWKGQCTSVTVSS	137	
Db	120	GVFDYWGQGTTLTVSS	135	

**RESULT 13**

US-09-748-960-6  
; Sequence 6, Application US/09748960

```

; Patent No. US20010046496A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Brettman, Lee R.
; APPLICANT: Fox, Judith A.
; APPLICANT: Allison, David Edward
;
; TITLE OF INVENTION: Method of Administering an Antibody
;
; FILE REFERENCE: 1855.2007-001
;
; CURRENT APPLICATION NUMBER: US/09/748,960
; CURRENT FILING DATE: 2000-12-27
;
; PRIOR APPLICATION NUMBER: US 09/550,082
; PRIOR FILING DATE: 2000-04-14
;
; NUMBER OF SEQ ID NOS: 14
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 6
;
; LENGTH: 180
;
; TYPE: PRT
;
; ORGANISM: Artificial Sequence
;
; FEATURE:
;
; OTHER INFORMATION: Portion of the heavy chain of LDP-0
;
; OTHER INFORMATION: chain signal peptide
;
; NAME/KEY: SITE
;
; LOCATION: (19)...(20)
;
; OTHER INFORMATION: signal peptide cleavage site
;
; US-09-748-960-6

```

	Query Match	66.1%	Score 539.5	DB 9	Length 180
	Best Local Similarity	65.2%	Pred. No. 1.1e-38		
	Matches 101	Conservative 18	Mismatches 33	Indels 3	Gaps 1
Qy	1	MECSWVFLFLSITTTGVHSGAYLQOSGAELVRSASVSKMCKASGYTLTISYNMHWVKTP	60		
Db	1	MKCTWVILFLVSTATSVHSGVQVLQSGAEVKKPGASVKVSKSGSYFTTSYMMHWVRQAP	60		
Qy	61	GOGLEWIGNTFPGNGDTYYNKKFKGKASLTADTSSSTAYMQISLTSSEDSAVVFCARG--	118		
Db	61	GORLEWIGEIDPSESNTNYYNKKFKGRVTLVDISASTAYMELSLRSEDATVYYCARGGY	120		
Qy	119	-NNEGALDYWGQGTSTVTSVSAKTTTPPPVPLVPGS	152		
Db	121	DGWDYATIDYWGQGTLVTVSSASKPGSPVFLPASP	155		

RESULT 14

```

US-09-269-921-105
; Sequence 105, Application US/09269921
; Publication No. US20030045691A1
; GENERAL INFORMATION:
; APPLICANT: Ono, Koichiro
; APPLICANT: Ohtomo, Toshihiko
; APPLICANT: Tsuchiya, Masayuki
; APPLICANT: Yoshimura, Yasushi
; APPLICANT: Kolshhara, Yasuo
; TITLE OF INVENTION: RESHAPED HUMAN ANTI-HM 1.24 ANTIBODY
; FILE REFERENCE: 35029-20007.00
; CURRENT APPLICATION NUMBER: US/09/269, 921
; CURRENT FILING DATE: 1999-04-01
; EARLIER APPLICATION NUMBER: PC1/JF97/03553
; EARLIER FILING DATE: 1997-10-03
; EARLIER APPLICATION NUMBER: JP 8-264756
; EARLIER FILING DATE: 1996-10-04
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
; LENGTH: 139
; TYPE: PRT
; ORGANISM: murine
; FEATURE:
; LOCATION:
; OTHER INFORMATION: Amino acid sequence of anti-HM 1.24
US-09-269-921-105

```

Query Match 65.0%; Score 530; DB 11; Length 139;  
Best Local Similarity 71.9%; Pred. No. 5.2e-38;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 22:02:45 ; Search time 18.4832 Seconds  
(without alignments)  
796.063 Million cell updates/sec

Title: US-08-836-455-4  
Perfect score: 816  
Sequence: 1 MECSWVFLFLSLITGVHQS.....TVSSAKTTPPPVPLVPGSL 153  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96158682 residues  
Total number of hits satisfying chosen parameters: 283308  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	568	69.6	475	2 S01321	Ig gamma-2b chain
2	563	69.0	469	2 S37483	Ig gamma-2a chain
3	560	68.6	474	1 G2M511	Ig gamma-2b chain
4	549	67.3	139	2 PS0024	Ig heavy chain pre
5	548.5	67.2	151	2 PL0011	Ig heavy chain pre
6	547	67.0	150	2 PN0444	Ig heavy chain v r
7	537.5	65.9	246	2 S38950	Ig gamma chain - m
8	537.5	65.9	446	2 S40295	Ig gamma-2a chain
9	533	65.3	178	2 S29594	Ig gamma chain (WM
10	529	64.8	137	2 H32513	Ig heavy chain pre
11	527.5	64.6	140	2 PH1482	Ig heavy chain v r
12	523.5	64.2	140	1 HVMSG7	Ig heavy chain pre
13	523.5	64.2	166	2 PL0012	Ig heavy chain pre
14	516.5	63.3	140	2 PH1489	Ig heavy chain v r
15	512	62.7	141	2 A39276	Ig heavy chain pre
16	510.5	62.6	138	2 E32513	Ig heavy chain pre
17	507.5	62.2	138	2 S21810	Ig heavy chain v r
18	507	62.1	133	2 PC1155	Ig heavy chain pre
19	506.5	62.1	140	2 PH1488	Ig heavy chain v r
20	502.5	61.6	135	2 PH1493	Ig heavy chain v r
21	502.5	61.6	140	2 PH1484	Ig heavy chain v r
22	502.5	61.6	140	2 PH1498	Ig heavy chain v r
23	501.5	61.5	135	2 PH1494	Ig heavy chain v r
24	499	61.2	135	2 PS0057	Ig heavy chain pre
25	497	60.9	139	2 A27609	Ig heavy chain pre
26	496.5	60.8	135	2 PH1492	Ig heavy chain v r
27	495	60.7	135	2 A30577	Ig heavy chain pre
28	494.5	60.6	140	2 PH1486	Ig heavy chain v r
29	493	60.4	139	1 MHMS18	Ig heavy chain pre

ALIGNMENTS

RESULT 1

S01321  
Ig gamma-2b chain precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 23-Jul-1999  
C:Accession: S01321  
Eur. J. Biochem. 176: 287-295, 1988  
A:Title: Expression in non-lymphoid cells of mouse recombinant immunoglobulin directed  
A:Reference number: S01320; MUID:88329081; PMID:3138116  
A:Accession: S01321  
A:Molecule type: mRNA  
A:Residues: 1-475 <DEL>  
A:Cross-references: EMBL:X13188; NID:g51780; PIDN:CAA31580.1; PID:g51781  
A:Note: this sequence was determined from the differentiated gene  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-475/Product: Ig gamma-2b chain #status predicted <MAT>  
F:159-223/Domain: immunoglobulin homology <IMM>

Query Match 69.6%; Score 568; DB 2; Length 475;  
Best Local Similarity 71.2%; Pred. No. 1.1e-37;  
Matches 109; Conservative 14; Mismatches 28; Indels 2; Gaps 1;  
Qy 1 MECSWVFLFLSLITGVHQSQAELVRSASVVKMSCKASGYTLTSYNMHWKQTP 60  
Db 1 MEWIFLFLSLITGVHQSQAELVRSASVVKMSCKASGYTLTSYNMHWKQRT 60  
Qy 61 GQGLEWIGNIFPGNGDTYYNQKPKASLTADTSSSTAYMQISSLTSEDSAVYFCARGNW 120  
Db 61 GQGLEWIGEYIPGSGNSYFNEKEKATLTVDKSSSTAYLHLSLTSEDSAVYFCAGPRQ 120  
Qy 121 EGAL--DYWGQGTSTVTVSSAKTTPPPVPLVPG 151  
Db 121 VGLLPFGYWGQGTSLVTAASAKTTPPSVYPLAPG 153

RESULT 2

S37483  
Ig gamma-2a chain - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999  
C:Accession: S37483  
R:Ducancel, F.F.D.  
submitted to the EMBL Data Library, February 1993  
A:Reference number: S37483  
A:Accession: S37483  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-469 <DUC>  
A:Cross-references: EMBL:X70423; NID:g406252; PIDN:CAA49868.1; PID:g406253

J. Biol. Chem. 269, 12345-12350, 1994

J. Biol. Chem. 269, 12345-12350, 1994

QY 1 MECSWVFLFLLSITTVGHSQAYLQQSGAELVRSGASVKMSCKASGYTLTSYNMHWVKQP 60

 $F; 1/\text{Mod}i$ 

ied site: pyrrolidone carboxylic acid (Gln) #status experimental





[illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 19:57:10 ; Search time 10.7819 Seconds  
(without alignments)  
667.331 Million cell updates/sec

Title: us-08-836-455-4  
Perfect score: 816  
Sequence: 1 MECSWVFLFLSITTCVHSG.....TVSSAKTTPPPVPLVPGSL 153

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Result No.	Score	Query Match	Length	DB	ID	Description
1	523.5	64.2	140	1	HV02_MOUSE	P01746 mus musculus
2	493	60.4	139	1	HV07_MOUSE	P01751 mus musculus
3	470	57.6	137	1	HV11_MOUSE	P01755 mus musculus
4	461	56.5	117	1	HV52_MOUSE	P06327 mus musculus
5	459.5	56.3	120	1	HV03_MOUSE	P01747 mus musculus
6	458	56.1	117	1	HV14_MOUSE	P01758 mus musculus
7	451	55.3	117	1	HV05_MOUSE	P01749 mus musculus
8	446	54.7	117	1	HV06_MOUSE	P01750 mus musculus
9	445.5	54.6	138	1	HV48_MOUSE	P03980 mus musculus
10	443	54.3	117	1	HV04_MOUSE	P01748 mus musculus
11	440	53.9	117	1	HV09_MOUSE	P01753 mus musculus
12	430	52.7	117	1	HV49_MOUSE	P06328 mus musculus
13	429.5	52.6	136	1	HV15_MOUSE	P01759 mus musculus
14	426.5	52.3	121	1	HV01_MOUSE	P01745 mus musculus
15	425.5	52.1	117	1	HV12_MOUSE	P01756 mus musculus
16	421.5	51.7	117	1	HV13_MOUSE	P01757 mus musculus
17	418	51.2	117	1	HV10_MOUSE	P01754 mus musculus
18	417	51.1	118	1	HV51_MOUSE	P06330 mus musculus
19	416	51.0	120	1	HV50_MOUSE	P06329 mus musculus
20	389	47.7	117	1	HV16_HUMAN	P23083 homo sapien
21	388	47.5	117	1	HV18_HUMAN	P01743 homo sapien
22	375	46.0	147	1	HVIC_HUMAN	P01744 homo sapien
23	342	41.9	114	1	HV00_MOUSE	P01741 mus musculus
24	333	40.8	136	1	HV16_MOUSE	P01783 mus musculus
25	321.5	39.4	142	1	HV01_RAT	P01805 rattus norv
26	308.5	37.8	121	1	HV30_HUMAN	P01771 homo sapien
27	308	37.7	122	1	HV36_HUMAN	P01768 homo sapien
28	304.5	37.3	144	1	HV43_MOUSE	P01819 mus musculus
29	301.5	36.9	117	1	HV1A_HUMAN	P01742 homo sapien
30	301	36.9	117	1	HV3C_HUMAN	P01764 homo sapien
31	300.5	36.8	117	1	HV41_MOUSE	P01811 mus musculus
32	299	36.6	119	1	HV40_MOUSE	P01810 mus musculus
33	299	36.6	126	1	HV3K_HUMAN	P01772 homo sapien

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	523.5	64.2	140	1	HV02_MOUSE	P01746 mus musculus
2	493	60.4	139	1	HV07_MOUSE	P01751 mus musculus
3	470	57.6	137	1	HV11_MOUSE	P01755 mus musculus
4	461	56.5	117	1	HV52_MOUSE	P06327 mus musculus
5	459.5	56.3	120	1	HV03_MOUSE	P01747 mus musculus
6	458	56.1	117	1	HV14_MOUSE	P01758 mus musculus
7	451	55.3	117	1	HV05_MOUSE	P01749 mus musculus
8	446	54.7	117	1	HV06_MOUSE	P01750 mus musculus
9	445.5	54.6	138	1	HV48_MOUSE	P03980 mus musculus
10	443	54.3	117	1	HV04_MOUSE	P01748 mus musculus
11	440	53.9	117	1	HV09_MOUSE	P01753 mus musculus
12	430	52.7	117	1	HV49_MOUSE	P06328 mus musculus
13	429.5	52.6	136	1	HV15_MOUSE	P01759 mus musculus
14	426.5	52.3	121	1	HV01_MOUSE	P01745 mus musculus
15	425.5	52.1	117	1	HV12_MOUSE	P01756 mus musculus
16	421.5	51.7	117	1	HV13_MOUSE	P01757 mus musculus
17	418	51.2	117	1	HV10_MOUSE	P01754 mus musculus
18	417	51.1	118	1	HV51_MOUSE	P06330 mus musculus
19	416	51.0	120	1	HV50_MOUSE	P06329 mus musculus
20	389	47.7	117	1	HV16_HUMAN	P23083 homo sapien
21	388	47.5	117	1	HV18_HUMAN	P01743 homo sapien
22	375	46.0	147	1	HVIC_HUMAN	P01744 homo sapien
23	342	41.9	114	1	HV00_MOUSE	P01741 mus musculus
24	333	40.8	136	1	HV16_MOUSE	P01783 mus musculus
25	321.5	39.4	142	1	HV01_RAT	P01805 rattus norv
26	308.5	37.8	121	1	HV30_HUMAN	P01771 homo sapien
27	308	37.7	122	1	HV36_HUMAN	P01768 homo sapien
28	304.5	37.3	144	1	HV43_MOUSE	P01819 mus musculus
29	301.5	36.9	117	1	HV1A_HUMAN	P01742 homo sapien
30	301	36.9	117	1	HV3C_HUMAN	P01764 homo sapien
31	300.5	36.8	117	1	HV41_MOUSE	P01811 mus musculus
32	299	36.6	119	1	HV40_MOUSE	P01810 mus musculus
33	299	36.6	126	1	HV3K_HUMAN	P01772 homo sapien

ALIGNMENTS

RESULT 1

HV02\_MOUSE STANDARD; PRT; 140 AA.

AC P01746;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Ig heavy chain V region 93G7 precursor.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RC SEQUENCE FROM N.A.

RP STRAIN=A/J;

RX MEDLINE=82152818; PubMed=6801765;

RA Sims J., Rabbitts T.H., Estess P., Slaughter C., Tucker P.W.,

RA Capra J.D.;

RT "Somatic mutation in genes for the variable portion of the immunoglobulin heavy chain.";

RL Science 216:309-311(1982).

CC -I- SIMILARITY: Contains 1 Immunoglobulin-like domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; J00493; AAA38128.1; -

DR PIR; A94264; HVMSG7.

DR HSSP; P01810; 2FBJ.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR003596; Ig\_v.

DR Pfam; PF00047; Ig; 1.

DR SMART; SM00406; IG; 1.

DR PROSITE; PS50835; IG\_LIKE; 1.

DR Immunoglobulin V region; Hybridoma; Signal.

FT SIGNAL 1 19

FT CHAIN 20 140 IG HEAVY CHAIN V REGION 93G7.

FT DOMAIN 20 139 IG-LIKE.

FT NON\_TER 140 140

SQ SEQUENCE 140 AA; 15514 MW; 25A4CBBE31DA5CE8 CRC64;

Query Match 64.2%; Score 523.5; DB 1; Length 140;

Best Local Similarity 72.1%; Pred. No. 1.9e-40;

Matches 101; Conservative 14; Mismatches 22; Indels 3; Gaps 1;

QY 1 MECSWVFLFLSITTCVHSGVAYLQSQAEALVRSGASVKMSCKASGYTLTSYNHHWVKQTP 60

Db 1 MGHSFIFLLSVTAGVHSEVQLQSQAEALVRSGSSVKMSCKASGYTFTSYGINWVKQRP 60

QY 61 GQGLEWIGNIFPGNGDTYYNQKPKGKASLTADTSSSTAYMQISLTSSEDSAVYFCARGNW 120

Db 61 GQGLEWICINPGNGYINNEKFKGKTLTVDKSSSTAYMQLRSLTSDSAVYFCARSHY 120  
 QY 121 EGA---LDYWGQGTSTVSS 137  
 Db 121 YGGSYDFYWGQGTPLTVSS 140

## RESULT 2

HV07\_MOUSE  
 ID HV07\_MOUSE STANDARD; PRT; 139 AA.  
 AC P01751; P01752;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ig heavy chain V region B1-8/186-2 precursor.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 EX MEDLINE=81234548; PubMed=6788376;  
 RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,  
 RA Baltimore D.;  
 RT "Heavy chain variable region contribution to the NPb family of  
 RT antibodies: somatic mutation evident in a gamma 2a variable region.";  
 RL Cell 24:625-637(1981).  
 CC -1- MISCELLANEOUS: THE B1-8 MU CHAIN MRNA WAS CLONED FROM A HYBRIDOMA  
 CC MAKING ANTIBODIES TO THE HAPTEN (4-HYDROXY-3-NITROPHENYL)ACETYL  
 CC (NPB ANTIBODIES).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: J00529; AAA38170.1; -  
 DR PIR: A90809; MMS18  
 DR PDB: 1A60; 27-MAY-98.  
 DR PDB: 1A6W; 15-JUL-98.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; Igv; 1.  
 DR PROSITE: PS00835; IG\_LIKE; 1.  
 DR Immunoglobulin V region; Signal; 3D-structure.  
 KW SIGNAL 1 19  
 FT CHAIN 20 139 IG HEAVY CHAIN V REGION B1-8/186-2.  
 FT DOMAIN 20 49 FRAMEWORK-1.  
 FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.  
 FT DOMAIN 55 68 FRAMEWORK-2.  
 FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.  
 FT DOMAIN 86 117 FRAMEWORK-3.  
 FT DOMAIN 118 124 D SEGMENT.  
 FT DOMAIN 125 139 JH2 SEGMENT.  
 FT DISULFID 41 115 BY SIMILARITY.  
 FT NON\_TER 139 139  
 SQ SEQUENCE 139 AA; 15419 MW; 1B57DD4FD0C9F465 CRC64;

Query Match 60.48; Score 493; DB 1; Length 139;  
 Best Local Similarity 69.18; Pred. No. 1e-37;  
 Matches 96; Conservative 15; Mismatches 26; Indels 2; Gaps 1;

QY 1 MECSWFLFLSLITGVHSGAYLQQSGAELVRSGASVKMSCKASGYTLTSYNMHWKQTP 60  
 Db 1 MGWSCIMFLAATATGVHSGVQLQPGAEVLPKASVKLSCKASGYTFTSYNMHWKQRP 60

QY 61 GQGLEWICINPGNGDYINQKFKKASLTADTSSSTAYMQLSSTSDSAVYFCARGNW 120  
 Db 61 GRGLEWIGRIDPNSGGTKYNEKPKSKATLTVDKFSSTAYMQLSSTSDSAVYCARVDY 120  
 QY 121 EGA---LDYWGQGTSTVSS 137  
 Db 121 YGGSYDFYWGQGTPLTVSS 139

## RESULT 3

HV11\_MOUSE  
 ID HV11\_MOUSE STANDARD; PRT; 137 AA.  
 AC P01755;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V region S43 precursor.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=81234548; PubMed=6788376;  
 RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,  
 RA Baltimore D.;  
 RT "Heavy chain variable region contribution to the NPb family of  
 RT antibodies: somatic mutation evident in a gamma 2a variable region.";  
 RL Cell 24:625-637(1981).  
 CC -1- MISCELLANEOUS: THE GAMMA-2A CHAIN MRNA WAS CLONED FROM A HYBRIDOMA  
 CC MAKING ANTIBODIES TO THE HAPTEN (4-HYDROXY-3-NITROPHENYL)ACETYL  
 CC (NPB ANTIBODIES).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: J00539; AAA38172.1; -  
 DR PIR: A02038; GZMS43.  
 DR HSP: P01810; 2FBJ.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; Igv; 1.  
 DR PROSITE: PS00835; IG\_LIKE; 1.  
 DR Immunoglobulin V region; Signal.  
 KW SIGNAL 1 19  
 FT CHAIN 20 137 IG HEAVY CHAIN V REGION S43.  
 FT DOMAIN 20 49 FRAMEWORK-1.  
 FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.  
 FT DOMAIN 55 68 FRAMEWORK-2.  
 FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.  
 FT DOMAIN 86 117 FRAMEWORK-3.  
 FT DOMAIN 118 122 D SEGMENT.  
 FT DOMAIN 123 137 JH2 SEGMENT.  
 FT DISULFID 41 115 BY SIMILARITY.  
 FT NON\_TER 137 137  
 SQ SEQUENCE 137 AA; 15200 MW; ADD5881BF44B8EC9 CRC64;

Query Match 57.6%; Score 470; DB 1; Length 137;  
 Best Local Similarity 66.48; Pred. No. 1.2e-35;  
 Matches 91; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

QY 1 MECSWFLFLSLITGVHSGAYLQQSGAELVRSGASVKMSCKASGYTLTSYNMHWKQTP 60  
 Db 1 MGWSCIMFLAATATGVHSGVQLQPGAEVLPKASVKLSCKASGYTFTSYNMHWKQRP 60

QY 61 GQGLEWICINPGNGDYINQKFKKASLTADTSSSTAYMQLSSTSDSAVYFCARGNW 120

```

Db 61 GRLEWIGRIDNSGGTYNEHFRSKATLTIDKPSSTAYMQLSSLTSEDSAVVYCYARYL 120
QY 121 EGALDYWGOGTSVTYSS 137
Db 121 GRFYDYWGOGTTLTVSS 137

RESULT 4
HV52_MOUSE
ID HV52_MOUSE STANDARD; PRT; 117 AA.
AC P06327;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region VH58 A1/A4 precursor.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85099340; PubMed=2578321;
RA Yancopoulos G.D., Alt F.W.;
RT "Developmentally controlled and tissue-specific expression of
unrearranged VH gene segments."
RL Cell 40:271-281(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
Db EMBL; M13787; AAA38499.1; -
DR PIR; A02029; HVMSAL.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION VH58 A1/A4.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 808C138856DFC9D CRC64;

Query Match 56.5%; Score 461; DB 1; Length 117;
Best Local Similarity 77.0%; Pred. No. 6.5e-35;
Matches 87; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 5 WYFLFLSLTTGVHQAQYVLSGAEVRSKASGYTLTSYNNHVMVKQTPGQGL 64
Db 5 WYFLFLSLTAGVHCPVQLQSQGPELVKPGALVKISKASGYTLTSYNNHVMVKQTPGQGL 64
QY 65 EWIGNLPNGGTYNKEKFGKASLTADTSSSTAYMQLSSLTSEDSAVVYFCAR 117
Db 65 EWIGWYPGDGSTKYNEKFGKATLTADKSSSTAYMQLSSLTSEDSAVVYFCAR 117

RESULT 5
HV03_MOUSE
ID HV03_MOUSE STANDARD; PRT; 120 AA.

```

```

AC P01747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region 36-65.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83131846; PubMed=6186498;
RA Siekevitz M., Geffer M.L., Brodeur P., Riblet R.,
RA Marshak-Rothstein A.;
RT "The genetic basis of antibody production: the dominant anti-arsonate
RT idiotypic response of the strain A mouse."
RL Eur. J. Immunol. 12:1023-1032(1982).
CC -!- MISCELLANEOUS: FROM ANALYSIS OF THE SIZES OF SEVERAL OTHER
CC DIFFERENTIATED GENES THAT HYBRIDIZE TO THIS ONE, THE AUTHORS
CC CONCLUDE THAT ALL OF THESE V REGIONS HAVE REARRANGED TO THE SAME J
CC SEGMENT, JH2.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR HSSP; P01789; IMCP.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR Immunoglobulin V region; Hybridoma.
FT DOMAIN 1 111 IG-LIKE.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13307 MW; FF04E4A167B654AF CRC64;

Query Match 56.3%; Score 459.5; DB 1; Length 120;
Best Local Similarity 75.4%; Pred. No. 9.1e-35;
Matches 89; Conservative 10; Mismatches 16; Indels 3; Gaps 1;

QY 23 LQSGAEVRSKASGYTLTSYNNHVMVKQTPGQGLWIGNIPFGNDTYNQK 82
Db 3 LQSGAEVRSKASGYTLTSYNNHVMVKQTPGQGLWIGNIPFGNDTYNQK 82
QY 83 FRGKASLTADTSSSTAYMQLSSLTSEDSAVVYFCARWEGA---LDYWGOGTSVTYSS 137
Db 63 FRGKTLTVDKSSSTAYMQLSSLTSEDSAVVYFCARSVYGGSYFYDYWGOGTTLTVSS 120

RESULT 6
HV14_MOUSE
ID HV14_MOUSE STANDARD; PRT; 117 AA.
AC P01758;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region 108A precursor.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81245215; PubMed=6789211;
RA Givol D., Zakut R., Rechavi G., Ram D., Cohen J.B.;
RT "Diversity of germ-line immunoglobulin VH genes."
RL Nature 292:426-430(1981).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC EMBL; J00488; AAA38519.1; -
CC -----
CC PIR; A02041; HVMS8A.
CC HSSP; P01810; 2FBJ.
CC MGI; 96486; Igh-VJ558.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003596; Ig_v.
CC Pfam; PF00047; Ig; 1.
CC SMART; SM00406; IGV; 1.
CC PROSITE; PS50835; IG-LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 108A.
FT DOMAIN 20 >117 IG-LIKE.
FT NON_TER 117
FT SEQUENCE 117 AA; 12972 MW; 428CB44DF25D1BC2 CRC64;
SQ
Query Match 56.1%; Score 458; DB 1; Length 117;
Best Local Similarity 74.4%; Pred. No. 1.2e-34;
Matches 87; Conservative 11; Mismatches 19; Indels 0; Gaps 0;
QY 1 MECSWVFLFLLSTTGVHQAAYLQSGAELVRGSAVKMSCKASGYTLTSYNNHHWYKQP 60
DB 1 MGWSVFLFLLSGTAGVSEVQLQSGPELVKPGASVKISCKASGYTFTDYNHHWYKQSH 60
QY 61 GQGLEWIGNIFPGNGDTYNNQKFKASLTADTSSTAYMOISLTSSEDSAVYFCAR 117
DB 61 GKLEWIGIYYPNGGTGYNQKFKSKATLTVDNSSSTAYMELSLTSSEDSAVYFCAR 117
RESULT 7
HV05_MOUSE STANDARD; PRT; 117 AA.
AC P01749;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ig heavy chain V region 3 precursor.
GN IGH-VJ558.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M.; Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPB family of
antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -1- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entitles requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 3.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117
FT SEQUENCE 117 AA; 13016 MW; 427C861C53975EDC CRC64;
SQ
Query Match 55.3%; Score 451; DB 1; Length 117;
Best Local Similarity 73.5%; Pred. No. 5.2e-34;
Matches 86; Conservative 12; Mismatches 19; Indels 0; Gaps 0;
QY 1 MECSWVFLFLLSTTGVHQAAYLQSGAELVRGSAVKMSCKASGYTLTSYNNHHWYKQP 60
DB 1 MGWSVFLFLLSGTAGVSEVQLQSGPELVKPGASVKISCKASGYTFTSYNDWYKQRP 60
QY 61 GQGLEWIGNIFPGNGDTYNNQKFKASLTADTSSTAYMOISLTSSEDSAVYFCAR 117
DB 61 GQGLEWIGNIYPSDSETHYNNQKFKATLTVDKSSSTAYMQLSLTSSEDSAVYFCAR 117
RESULT 8
HV06_MOUSE STANDARD; PRT; 117 AA.
AC P01750;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 102 precursor.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M.; Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPB family of
antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981)
CC -1- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.
CC -----
CC PIR; A02032; HVMS02.
CC HSSP; P01810; 2FBJ.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003596; Ig_v.
CC Pfam; PF00047; Ig; 1.
CC SMART; SM00406; IGV; 1.
CC PROSITE; PS50835; IG-LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 102.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117
FT SEQUENCE 117 AA; 12867 MW; 740A65DD851FCA8C CRC64;
SQ
Query Match 54.7%; Score 446; DB 1; Length 117;
Best Local Similarity 74.1%; Pred. No. 1.5e-33;
```

ID	HV04_MOUSE	STANDARD;	PRT;	117 AA.
----	------------	-----------	------	---------

RI ANTIBODIES: SOMATIC MUTATION EVIDENT IN A GAMMA 2A VARIABLE REGION.";





[illegible]

RESULT 14
HV01_MOUSE
ID HV01_MOUSE STANDARD; PRT; 121 AA.
P01745;
DT DT 21-JUL-1986 (Rel. 01, Created)
DT DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE DE Ig heavy chain V region MPC 11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81053741; PubMed=6253904;
RA Zakut R., Cohen J.; Givol D.;
RT "Cloning and sequence of the cDNA corresponding to the variable
RL region of immunoglobulin heavy chain MPC11.";
RN Nucleic Acids Res. 8:3591-3601(1980).
[2]
RP REVISIONS.
RP Zakut R., Cohen J., Givol D.;
RA Nucleic Acids Res. 8:4839-4840(1980).
RL -! MISCELLANEOUS: THIS SEQUENCE WAS TRANSLATED FROM AN MRNA ISOLATED
CC FROM A MYELOMA THAT SECRETES IGB2B.
CC -! SIMILARITY: Contains 1 Immunoglobulin-like domain.
CC FIR; A93708; GVMSJ1.
DR HSSP; P01810; 2FBJ.
DR InterPro: IPR007110; Ig-Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 112 IG-LIKE.
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 113135 MW; 227AEFF3EC56ED0BF CRC64;

**RESULT 15**

```

HV12_MOUSE
ID HV12_MOUSE STANDARD; PRT; 117 AA.
AC P01756;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region MOPC 104E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE ASN-55.
MEDLINE=8075344; PubMed=6816276;
RA Kehry M.R., Fuhrman J.S., Schilling J.W., Rogers J., Sibley C.H.,
RT Hood L.E.;
RC "Complete amino acid sequence of a mouse mu chain: homology among
RT heavy chain constant region domains.";
RL Biochemistry 21:5415-5424(1982).
CC -I- MISCELLANEOUS: THE SEQUENCE OF THE LIGHT CHAIN OF THIS IGM MYELOMA
CC PROTEIN HAS ALSO BEEN DETERMINED.
CC -I- MISCELLANEOUS: THIS PROTEIN BINDS DEXTRAN.
CC -I- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02039; MEMS4E.
DR HSSP; P01789; LMCP.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin V region; Glycoprotein.
FT DOMAIN 1 116
IG-LIKE.
FT DISULFID 22 96
BY SIMILARITY.
FT CARBOHYD 55 55
N-LINKED (GLCNAC... ) (COMPLEX).
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12983 MW; 3CF8ACE4BE447E41 CRC64;

Query Match 52.1%; Score 425.5; DB 1; Length 117;.
Best Local Similarity 70.3%; Pred. No. 1e-31;
Matches 83; Conservative 13; Mismatches 21; Indels 1; Gaps 1;

QY 20 QAYLQSGAELVRSASVKMSCKASYTTLTSYNHHVVKQTGGGLEWIGNIFPGNGTTY 79
db : : ||||| ||| : ||||| ||||| | : ||||| : ||||| : ||||| : |||||
1 EVQLQGSPGLVPKGASVKMSCKASYTTFDYYMKWKVKQSHGSKSLSEWIGDPNNNGTSY 60

QY 80 NQKFCKASLTADTSSSTAYMQLSSITSDSAVFYFCAGNWEGALDWGGGTSTVTSS 137
db ||| ||||| : ||| : ||||| ||||| ||||| ||||| : ||| : ||||| |||||
61 NORFKRALTVDKSSSTAYMQNLSTSDSAVIYCARI-DIWDFYFWGAGTTVTSS 117

Search completed: August 30, 2003, 22:04:35.
Job time : 11.7819 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2003, 21:05:00 ; Search time 36.453 Seconds  
(without alignments)  
1083.094 Million cell updates/sec

Title: US-08-836-455-4  
Perfect score: 816  
Sequence: 1 MCSWVFLFLSITGVHSQ.....TVSSAKTTPPPVPLVPGSL 153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriapi:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	589.5	72.2	473	11 Q9D8L4	Q9D8L4 mus musculus
2	581	71.2	463	11 Q99LC4	Q99LC4 mus musculus
3	569	69.7	473	11 Q99L25	Q99L25 mus musculus
4	549.5	67.3	168	11 Q8VDC9	Q8VDC9 mus musculus
5	544	66.7	613	11 Q8VXC7	Q8VXC7 mus musculus
6	526.5	64.5	481	11 Q91WT1	Q91WT1 mus musculus
7	524.5	64.3	474	11 Q8R3H6	Q8R3H6 mus musculus
8	521	63.8	278	11 Q921K1	Q921K1 mus musculus
9	521	63.8	482	11 Q8K172	Q8K172 mus musculus
10	516.5	63.3	489	11 Q8VXC4	Q8VXC4 mus musculus
11	515.5	63.2	468	11 Q99L31	Q99L31 mus musculus
12	506.5	62.1	142	11 Q924Q1	Q924Q1 mus musculus
13	499	61.2	480	11 Q8K0Z4	Q8K0Z4 mus musculus
14	494.5	60.6	481	11 Q8VVC5	Q8VVC5 mus musculus
15	493	60.4	488	11 Q91WR1	Q91WR1 mus musculus
16	489	59.9	488	11 Q8K0F2	Q8K0F2 mus musculus

17	482	59.1	143	11 Q924R0	Q924R0 mus musculus
18	479.5	58.8	481	11 Q91WT3	Q91WT3 mus musculus
19	478	58.6	145	11 Q924R4	Q924R4 mus musculus
20	474	58.1	145	11 Q924P7	Q924P7 mus musculus
21	474	58.1	145	11 Q924R1	Q924R1 mus musculus
22	473.5	58.0	140	11 Q924P8	Q924P8 mus musculus
23	473.5	58.0	146	11 Q924Q3	Q924Q3 mus musculus
24	472	57.8	145	11 Q924Q6	Q924Q6 mus musculus
25	471	57.7	143	11 Q924P9	Q924P9 mus musculus
26	469	57.5	145	11 Q924Q9	Q924Q9 mus musculus
27	467	57.2	145	11 Q924Q7	Q924Q7 mus musculus
28	466.5	57.2	142	11 Q924Q2	Q924Q2 mus musculus
29	466.5	57.2	146	11 Q924R8	Q924R8 mus musculus
30	465	57.0	143	11 Q924Q0	Q924Q0 mus musculus
31	464	56.9	143	11 Q924R7	Q924R7 mus musculus
32	464	56.9	143	11 Q924P6	Q924P6 mus musculus
33	464	56.9	145	11 Q924R3	Q924R3 mus musculus
34	463.5	56.8	140	11 Q924R2	Q924R2 mus musculus
35	463	56.7	137	11 Q924R6	Q924R6 mus musculus
36	459	56.2	143	11 Q91VA2	Q91VA2 mus musculus
37	458	56.1	143	11 Q91V67	Q91V67 mus musculus
38	453	55.5	141	11 Q924Q4	Q924Q4 mus musculus
39	451.5	55.3	497	4 Q8WY24	Q8WY24 homo sapien
40	450	55.1	118	11 Q921C4	Q921C4 mus musculus
41	449	55.0	143	11 Q924Q5	Q924Q5 mus musculus
42	447.5	54.8	146	11 Q924Q8	Q924Q8 mus musculus
43	447.5	54.8	614	4 Q96GA6	Q96GA6 homo sapien
44	446.5	54.7	117	11 Q9QXE9	Q9QXE9 mus musculus
45	445.5	54.6	144	11 Q924P5	Q924P5 mus musculus

ALIGNMENTS

RESULT 1

Q9D8L4	PRELIMINARY;	PRT;	473 AA.
ID	Q9D8L4		
AC	Q9D8L4;		
DT	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	1810060009RIK protein.		
GN	IGH-1 OR 1810060009RIK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Pancreas;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Famanaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadoita K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,		
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,		
RT	Hayashizaki Y.;		
RL	"Functional annotation of a full-length mouse cDNA collection.";		
RL	Nature 409:685-690(2001).		
DR	EMBL; AK007918; BAB25349.1; -.		
DR	HSSP; P01842; 7FAB.		



```
DR SMART: SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 168
SQ SEQUENCE 168 AA; 18293 MW; 1E3719FCC0E72723 CRC64;

Query Match 67.3%; Score 549.5; DB 11; Length 168;
Best Local Similarity 69.5%; Pred. No. 3e-44;
Matches 105; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

QY 1 MECSWVFLFLSITTVGHVSQAYLQOQSGAELVRSGASVKMSCKASGYTLTSYNMHWKQTP 60
DB 11 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 MEWTWVFLFLSVTAGVHVSQVQLQOQSGAELMKPGASVKISCKATGYTFSSYIWDWKQRP 60
QY 61 GQGLEWIGNFPNGDYYNQKFKGASLTADTSSSTAYMOISLTSSEDSAVYFCAR-GN 119
DB 11 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 61 GHGLEWIGELPGSGRTNYNEKFKGKATFTADTSSNTAYIQFSLTSSEDSAVYCYCARRLG 120
QY 120 WEGALDYWGQGTSVTVSSAKTTPPPVPLV 150
DB 11 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 121 SRWYFDVWGAGTTVTVSSTKTTAPSVYPLAP 151

RESULT 5
Q8VCX7 PRELIMINARY; PRT; 613 AA.
AC Q8VCX7;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical 67.9 kDa protein.
GN IGH-6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC018315; AAH18315.1;
DR MGD; MGI:96448; Igh-6.
DR InterPro; IPR007110; Igh-like.
DR InterPro; IPR003006; Igh_MHC.
DR Pfam; PF00047; Igh_V.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 5.
DR PROSITE; PS00290; IG_MHC; 3.
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 67855 MW; 41A9384DD4C22862 CRC64;

Query Match 66.7%; Score 544; DB 11; Length 613;
Best Local Similarity 69.5%; Pred. No. 5.1e-43;
Matches 105; Conservative 17; Mismatches 25; Indels 4; Gaps 2;

QY 1 MECSWVFLFLSITTVGHVSQAYLQOQSGAELVRSGASVKMSCKASGYTLTSYNMHWKQTP 60
DB 11 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 MEWTWVFLFLSVTAGVHVSQVQLQOQSGAELMKPGASVKISCKATGYTFSSYIWDWKQRP 60
QY 61 GQGLEWIGNFPNGDYYNQKFKGASLTADTSSSTAYMOISLTSSEDSAVYFCAR--G 118
DB 11 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 61 GHGLEWIGELPGSGRTNYNEKFKGKATFTADTSSNTAYMQLSLSSEDSAVYCYCARRLG 120
QY 119 WEGALDYWGQGTSVTVSSAKTTPPPVPLV 149
DB 11 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 121 RW--YFDVWGAGTTVTVSSESQSPNVPLV 149

RESULT 6
Q91WT1 PRELIMINARY; PRT; 481 AA.
ID Q91WT1
AC Q91WT1;
```

```
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical 52.1 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013490; AAH13490.1;
DR InterPro; IPR007110; Igh-like.
DR InterPro; IPR003006; Igh_MHC.
DR InterPro; IPR003596; Igh_V.
DR Pfam; PF00047; Igh_4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 2.
KW Hypothetical protein.
SQ SEQUENCE 481 AA; 52105 MW; 97DF68D159463F65 CRC64;

Query Match 64.5%; Score 526.5; DB 11; Length 481;
Best Local Similarity 70.3%; Pred. No. 1.7e-41;
Matches 102; Conservative 13; Mismatches 27; Indels 3; Gaps 2;

QY 5 WVFLLSITTVGHVSQAYLQOQSGAELVRSGASVKMSCKASGYTLTSYNMHWKQTPGQGL 64
DB 11 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 5 WIFLLSGTAGVOCVOLLOSGPELVKPGASVKISCKASGYTFTSYIIHWKORPGQGL 64
QY 65 EWIGNFPNGDYYNQKFKGASLTADTSSSTAYMOISLTSSEDSAVYFCAR-GNWEGA 123
DB 11 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 65 VWIGWIYPGDNTKYNEKFKGKATFTADTSSSTAYMFLSLSSEDSAVYFCARGGW--A 122
QY 124 LDYWGQGTSTVSSAKTTPPPVPL 148
DB 11 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 123 FDYWGQGTTLTVSSEPAEPTIYPL 147

RESULT 7
Q8R3H6 PRELIMINARY; PRT; 474 AA.
ID Q8R3H6
AC Q8R3H6;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical 51.7 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025447; AAH25447.1;
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR007110; Igh-like.
DR InterPro; IPR003006; Igh_MHC.
DR InterPro; IPR003596; Igh_V.
DR Pfam; PF00047; Igh_3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 474 AA; 51748 MW; 8608B57C6CD2874A CRC64;

Query Match 64.3%; Score 524.5; DB 11; Length 474;
Best Local Similarity 66.4%; Pred. No. 2.6e-41;
Matches 101; Conservative 20; Mismatches 30; Indels 1; Gaps 1;
```



```
Db 121 RGNVDGLAWFVWGQGLTVVSAEPAREPTIYPL 155
RESULT 11
Q99L31
ID Q99L31 PRELIMINARY; PRT; 468 AA.
AC Q99L31;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to RIKEN cDNA 1810060C09 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AAH03878.1; -
DR HSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 468 AA; 51661 MW; 96352328B332ADB CRC64;

Query Match 63.2%; Score 515.5; DB 11; Length 468;
Best Local Similarity 64.2%; Pred. No. 1.8e-40;
Matches 97; Conservative 23; Mismatches 30; Indels 1; Gaps 1;

QY 1 MEGSWIFLLSITTVGHSQAYLQSGAELVRSGASVKMSCKASGYTLTSYNNHWKQTP 60
DB 1 MKCSWIFFLMAVIGVNSEVQQLQSGAELVRSGASVKLSCTAGFNKDSLMHWKQRP 60
QY 61 GQLEWIGNIFPGNGDTYYNQKFKGASLTADTSSSTAYMQISLTSSEDSAVYFCARG-N 119
DB 61 EQLEWIGWIDPEDGETKYPKQDKATITADTSSNTAYLQSLTSSEDTAYIYCARNLL 120
QY 120 WEGALDWGQGTSTVSSAKTTPPPVPLVP 150
DB 121 YGGYDWGQGTITVSSAKTTPAPSVYPLAP 151

RESULT 12
Q924Q1
ID Q924Q1 PRELIMINARY; PRT; 142 AA.
AC Q924Q1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE V23-D-J-C mu protein (fragment).
GN V23-D-J-C MU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A.
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB069913; BAB63929.1; -
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.

Query Match 61.2%; Score 499; DB 11; Length 480;
Best Local Similarity 65.3%; Pred. No. 6.8e-39;
Matches 98; Conservative 18; Mismatches 32; Indels 2; Gaps 2;

QY 5 WVFLLSITTVGHSQAYLQSGAELVRSGASVKMSCKASGYTLTSYNNHWKQTPGQGL 64
DB 5 WIFLLSGTAGVHCQVQLQSGPELVKPGALVKISCKASYTFTSFDSIMWKQPGQGP 64
QY 65 EWIGNIFPGNGDTYYNQKFKGASLTADTSSSTAYMQISLTSSEDSAVYFCARGNWEGAL 124
DB 65 EWIGWISFGDGSSEYNEKFKGKATLTADKSSNTAYMHLSSLTSENSAVYFCARSKL-GGF 123
QY 125 DWGQGTSTVSSAKTTPPPVPL-VPGSL 153
DB 124 AYWGQGTTLTVSAESARNPTIYPLTLPAL 153

RESULT 14
Q8VCV5
ID Q8VCV5 PRELIMINARY; PRT; 480 AA.
AC Q8VCV5;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to expressed sequence AI893585.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC029188; AAH29188.1; -
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 3.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 2.
SQ SEQUENCE 480 AA; 51645 MW; 8690A63C669CDBED CRC64;

Query Match 61.2%; Score 499; DB 11; Length 480;
Best Local Similarity 65.3%; Pred. No. 6.8e-39;
Matches 98; Conservative 18; Mismatches 32; Indels 2; Gaps 2;

QY 5 WVFLLSITTVGHSQAYLQSGAELVRSGASVKMSCKASGYTLTSYNNHWKQTPGQGL 64
DB 5 WIFLLSGTAGVHCQVQLQSGPELVKPGALVKISCKASYTFTSFDSIMWKQPGQGP 64
QY 65 EWIGNIFPGNGDTYYNQKFKGASLTADTSSSTAYMQISLTSSEDSAVYFCARGNWEGAL 124
DB 65 EWIGWISFGDGSSEYNEKFKGKATLTADKSSNTAYMHLSSLTSENSAVYFCARSKL-GGF 123
QY 125 DWGQGTSTVSSAKTTPPPVPL-VPGSL 153
DB 124 AYWGQGTTLTVSAESARNPTIYPLTLPAL 153

RESULT 14
Q8VCV5
```

```

ID Q8VCV5 PRELIMINARY; PRT; 481 AA.
AC Q8VCV5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 52.3 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast tumor;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC018455; AAH18455.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 2.
KW Hypothetical protein.
SQ SEQUENCE 481 AA; 52326 MW; 52B44C5826807143 CRC64;

Query Match 60.6%; Score 494.5; DB 11; Length 481;
Best Local Similarity 63.2%; Pred. No. 1.8e-38;
Matches 98; Conservative 21; Mismatches 33; Indels 3; Gaps 3;

Qy 1 MECSWVFLLSITGVHSGAYLQOAGELVRSGASVKMSCKASGYTLTSYNMHWKQTP 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MEWSWIFLLSGTGVHSEIQLQOAGPELVKPGTSVKVSKASGYTFIDYNIYWKOSH 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 GQGLEWIGNIPPGNDTYYNOKFKGKASLTADTSSSTAYMOISSITSDSAVYFCARGN 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GKSLWIGIDINPYNGGTSYNOKFKGKATLTVDKSSIAYMQLNNLTSDSAFYTCAR-EW 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 EGA-LDYWGQGTSTVSSAKTTPPPVYPL-VPGLS 153
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 YGANFAFWGQGLTVTVSAESARNPTIYPLTLPPL 154
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 15
Q91WR1 PRELIMINARY; PRT; 488 AA.
ID Q91WR1
AC Q91WR1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 53.0 kDa protein.
GN IGH-VJ558 OR AI893585.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013539; AAH13539.1; -.
DR MGD; MGI:96486; Igh-VJ558.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 2.
KW Hypothetical protein.
SQ SEQUENCE 488 AA; 52964 MW; F12068460B400B9D CRC64;
```

```

Query Match 60.4%; Score 493; DB 11; Length 488;
Best Local Similarity 61.3%; Pred. No. 2.6e-38;
Matches 95; Conservative 24; Mismatches 28; Indels 8; Gaps 2;

Qy 1 MECSWVFLLSITGVHSGAYLQOAGELVRSGASVKMSCKASGYTLTSYNMHWKQTP 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MGWNWIFLLSGTAGIYSEVQLQOAGPELVKPGASVKLSCKASGYTIDYIYVWKOSH 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 GQGLEWIGNIPPGNDTYYNOKFKGKASLTADTSSSTAYMOISSITSDSAVYFCARG-- 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GKSLWIGIDINPYNGGTSYNOKFKGKATLTVDKSSIAYMQLNNLTSDSAVYFCARGPV 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 119 -----NWEGALDYWGQGTSTVTVSSAKTTPPPVYPL 148
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 YYSYFSYDRG-DYWGQGLTVTVSAEPAREPTIYPL 154
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Search completed: August 30, 2003, 22:06:01
Job time : 38.453 secs
```